DISSERTATION

Design and Evaluation of Photonic Neural Network for Hardware-Accelerated Artificial Intelligence Applications

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der technischen Wissenschaften

unter der Betreuung von

Univ.Prof. Dipl.-Ing. Dr.techn. Michael Waltl
and
Dr. Bernhard Schrenk

eingereicht an der Technischen Universität Wien
Fakultät für Elektrotechnik und Informationstechnik
von

Margareta Vania Stephanie, MSc.
Matrikelnummer: 12112242

Wien, im Oktober 2024

Keep Going

“When things go wrong, as they sometimes will,
When the road you’re trudging seems all uphill,
When the funds are low and the debts are high,
And you want to smile, but you have to sigh,
When care is pressing you down a bit,
Rest, if you must, but don’t you quit.

Life is queer with its twists and turns,
As every one of us sometimes learns,
And many a failure turns about,
When he might have won had he stuck it out;
Don’t give up though the pace seems slow-
You may succeed with another blow.

Often the goal is nearer than,
It seems to a faint and faltering man,
Often the struggler has given up,
When he might have captured the victor’s cup,
And he learned too late when the night slipped down,
How close he was to the golden crown.

Success is failure turned inside out-
The silver tint of the clouds of doubt,
And you never can tell how close you are,
It may be near when it seems so far,
So stick to the fight when you’re hardest hit-
It’s when things seem worst that you must not quit”

by Edgar Guest, “Keep Going”, published in Detroit Free Press (4 March 1921)

Acknowledgement

Doing a PhD is an experience that no one can understand until one experiences it for themselves. I never thought I would do it, but here I am, writing some kind of epilogue to my life journey.

First, I want to express my sincere gratitude to my supervisor, Dr. Bernhard Schrenk from AIT Austrian Institute of Technology, who has mentored me for these three years. I truly appreciate all the time and effort you have invested in guiding me. Second, I cannot thank enough my TU Wien supervisors, Prof. Michael Waltl and Prof. Tibor Grasser. I am incredibly grateful for your mentorship and support. The time and effort you want to spend on my PhD supervising are outstanding. Furthermore, I want to thank my reviewers, Prof. Nikos Pleros and Ass. Prof. Ripalta (Patty) Stabile for their willingness to be part of my PhD journey. I am glad that I met you at ECOC 2024.

Additionally, I am grateful for the opportunity to be part of AIT because I have learned a lot, submitted publications, and presented at many conferences worldwide. Thank you to the Austrian Research Promotion Agency FFG, which has funded the JOLLYBEE project as the main research topic for my thesis. I also want to express my appreciation to Dr. Lam Pham and Dr. Alexander Schindler, our collaborators in the JOLLYBEE Project, who provided a multilayer perception neural network model in Matlab and Python for Iris flower dataset. Thank you for your help and support on this project.

Next, I want to thank Aina Val-Marti for being a supportive colleague during this difficult time. Your words of encouragement and acts of kindness have helped me through a difficult time in my life, and I appreciate your help more than words can say. We can do this because we do it together. Thanks to Florian Honz and Florian Kutschera for their sincere help in the lab and for a good time in the office. Special thanks to other photonics members at AIT: Nemanja Vokić, Dinka Milovančev, and Winfried Boxleitner, for their help, motivation, and support. Your encouragement has been a source of strength, and I am grateful for your advice and guidance. Furthermore, thank you to Markus Hofer, Anja Dakić, Hamed Radpour, Pascal Thiele, and Mariana Ferreira-Ramos for the entertaining lunchtime and coffee break. In addition, I want to thank all the wireless and quantum members at AIT. I cannot mention the names individually, but thank you for sharing your knowledge and experience with me. I will never forget to thank Irene Reinisch, Birgit Lechner, Elisabeth Seifert, and Manuela Kos, who have always helped me with all the required administrative work at AIT.

I would like to give the biggest appreciation to my fiancée, Philip Rafetseder. Your unwavering support and love have been my anchor during the stormy seas of my PhD journey. Your tenderness and compassion have made the difficult times easier to bear, and I am truly fortunate to have you in my life. Additionally, I cannot thank enough my best friends, Bayu Rakhmatullah and David Senjaya. We keep this long-distance friendship for more than 10 years and I guess it will keep go on. Thank you for walking through the difficult path with me. I also would like to give thanks to my family: my deceased father (Antonius Hernas Setiady), my mom (Herlianti), my brother (Davin Ardian), and my sister-in-law (Esterin Violety Angelina). Thank you for all the encouragement and especially for letting me go far away from home without asking too many questions about my decision. It seems like the path that I take is getting more rough and not easy. However, thank you for believing in me.

In addition, I cannot forget my daily dose of sunshine in Austria. The Crete family: Aina, Sonu, Amr, and Rosa, thanks for the lovely time we spent together, either hiking, traveling, or tea time together. I appreciate your support and concern more than words can express. Thank you for being there for me. Thank you to NÖ-card geng: Lisa, Yuda, Gavrila, and Iqbal. I had a fantastic experience traveling with you every weekend and finding hidden gems in Austria. I also want to express my heartfelt thanks to my beautiful friend, Ani. Your bright and positive attitude always cheers me up. Thank you for your concern and unwavering support. Athina, I will also never forget you. You were next to my office the first time I arrived at AIT. Your kindness and support have been a great help to me. Thank you for everything, including your Pilea. It grows so much now in the office. I am also truly grateful to my Hockegasse Dinner family: Paulina, Yuga, Marina, Ryan, Marie, Radit, and Larissa. I had an amazing time of cooking and get-together. Thank you for lending listening ears or offering helping hands. I feel so lucky to have found you in Austria. Eventually, thank you to the KKIA (Indonesische Christliche Gemeinde) family. A warm atmosphere makes me feel at home in Indonesia through our conversation and overflowing Indonesian cuisine. I cannot mention the names one by one, but because of you all, I do not feel lonely in this beautiful country of Austria.

Abstract

Nearly two-thirds of the global population now have internet access. As a result, annual global Internet traffic has increased exponentially to 4.8 Zettabytes (ZB) by 2022. The rise of users has driven the growth of data centers and smart devices for Internet-of-Things (IoT), expected to take a share of 23% of the global energy footprint by 2030. This problem is going mutually with the von-Neumann computing architecture bottleneck. The saturation of clock frequency in the central processing unit (CPU) has flattened out the current computing machine efficiency, which cannot sustain future data processing and generation demand. Eventually, the digital brick wall leads to a computing gap where a new computing technology is necessary.

On the other hand, the brain is a powerful natural computer within humans, which has a large scale and an efficient algorithm beyond the human-built computer. The brain has the capability to perform $10^{13} - 10^{16}$ operations per second with a power of only 15-25 W, which is eight orders of magnitude more efficient than the current digital electronic supercomputer. The parallel processing of the brain's neural network is the core of high-speed computation, which has been studied to develop artificial intelligence (AI). The success of AI, which has empowered decision-making machines, has pushed the investigation of brain-inspired computers at the physical level or a so-called neuromorphic computing.

Photonics, with its inherent parallelism and proven ability to offer low latency and fast signal processing in the GHz range, is a promising platform for neuromorphic computing. This potential of photonics for neuromorphic computing offers hope for the future of technology, with the possibility of developing faster and more efficient computing systems.

The neural network is based on two essential ingredients: the linear weighted summation, where all the neural inputs are weighted and summed, followed by the nonlinear activation function for the decision boundary. One approach to performing analog neuromorphic processing is to employ it in the optical domain, where the equivalent neural network functions have to be identified. Towards this direction, this thesis plays an important role in finding the key ingredients in the photonics realm based on the frequency coding of the synaptic signals.

The first part presents a synaptic receptor that functionally integrates the weighting and summation functions. The matrix multiplication based on the frequency coding of synaptic signals can be performed by employing the wide bandwidth of a semiconductor optical amplifier (SOA) and an electroabsorption modulator (EAM). It will be shown that two spike trains can be simultaneously processed with opposite polarities and detected as the weighted sum. The result proves that the synaptic receptor can carry multiple colors of synapses and can potentially scale the optical neural network (ONN).

Afterwards, this thesis investigates the photonic activation function, focusing on the rectified linear unit (ReLU) function. The ReLU unit is performed by optical frequency coding of signals via a linear transmission of an optical filter and a subsequent photodetector with its noise for clipping the signal. A digital neural network (DNN) based on the Iris flower classification case is used as a benchmark. It will be demonstrated that a small penalty of 3% in accuracy is achieved when transferring the classification challenge from the digital to the optical domain. This renders that the scheme can be efficiently employed in the ONN.

Next, the constituent ingredients are combined to perform the photonic neuron. Multiple neural sub-circuits can be collapsed over the photonic ReLU. The demonstration reveals that the weighting functionality can be integrated over the same ReLU unit, which motivates the simplification and cost-optimization of ONN scaling. A low penalty of 1% can be achieved in terms of accuracy compared to DNN.

Nevertheless, optoelectronic conversion is the main bottleneck between the linear and nonlinear operations in the photonic neuron. A translucent concatenation scheme is discussed afterwards, where the photocurrent of the weighted sum directly drives the subsequent FM translator in the activation circuit without the need for electronic amplification. This investigation promises an energy-efficient interface that can be scaled to a smaller footprint.

Eventually, the functional complete neuron in neuromorphic processing is presented in a space-switch architecture. Especially, the space-switch architecture can be made reconfigurable as a so-called photonic processing unit (ΦPU) to employ other signal processing functions, taking advantage of re-using the optical elements. It will be confirmed that ΦPU has a penalty-free operation. Furthermore, the neuromorphic processing in the space-switch is compared with the one in wavelength-routed architecture by employing the translucent concatenation scheme. A penalty <2% can be achieved for multiple optical/digital sub-circuits implementation scenarios.

The novelties presented in the thesis contribute to the neuromorphic photonics field and the possibility of scaling the ONN in a smaller device footprint.

Contents

List of Figures

List of Tables

List of Abbreviations

$L-I$	light-current
ΦPU	photonic processing unit
AI	artificial intelligence
ANN	artificial neural network
AOTF	acousto-optic tunable filter
APD	avalanche photodiode
ASE	amplifier spontaneous emission
AWG	arbitrary waveform generator
AWGs	arrayed waveguide gratings
BER	bit error ratio
CML	chirp-managed modulated laser
CMOS	complementary metal-oxide semiconductor
CO_2	carbon dioxide
CPU	central processing unit
CS-DSB	carrier-suppressed double-sideband
CWDM	coarse-wavelength division multiplexing
DBR	distributed Bragg reflector
DC	direct current
DFB	distributed feedback
DI	delay interferometer
DML	directly modulated laser
DNN	digital NN
DSP	Digital signal processing
e/o	electrical-to-optical
EA	electroabsorption
EAM	electroabsorption modulator
EDFA	erbium-doped fiber amplifier
EML	electroabsorption-modulated laser
ER	extinction ratio
FBGF	fiber bragg grating filter

FDBR	front DBR
FM	frequency-modulated
FSR	free spectral range
FWHM	full width at half maximum
Ge/Si	Germanium/Silicon
GPUs	graphic processing units
HiCANN	High-Count Analogue Neural Network
ICT	information communication technology
IF	intermediate frequency
IL	interleaver
IM	intensity-modulated
IM/DD	intensity modulation with direct detection
InP	Indium Phosphite
IQM	inphase/quadrature modulator
LIDAR	light detection and ranging
LNA	low-noise amplifier
LO	laser oscillator
LSB	least significant bit
MAC	multiply accumulate
MEMS	micro electro mechanical systems
MLP	multilayer perceptron
MPW	multi project wafer
MQW	multi-quantum well
MRR	micro-ring resonator
MSB	most significant bit
MZI	Mach-Zehnder interferometer
MZM	Mach-Zehnder modulator
NEP	noise equivalent power
NN	neural network
NRZ	non-return-to-zero
NV	nitrogen-vacancy
o/e	optical-to-electrical
o/e/o	optical-to-electrical-to-optical
ONN	Optical neural network
OOK	on-off-keying
OSR	optical spectrum reshaper
OTF	optical tunable filter
PAM4	pulse amplitude modulation with four levels
PC	polarization controller

PDK	process design kit
PIC	photonic integrated circuit
PIN	positive-intrinsic-negative diode
QCSE	quantum-confined stark effect
qubits	quantum bits
R&D	research and development
RC	reservoir computing
RDBR	rear DBR
ReLU	rectified linear unit
RF	radio frequency
RNN	recurrent neural network
ROP	received optical power
Rx	receivers
SDM	Space-division multiplexing
SMF	single-mode fiber
SNR	signal-to-noise ratio
SOA	semiconductor optical amplifier
SOA-REAM	SOA-reflective EAM
SoI	silicon-on-insulator
TDM	Time-division multiplexing
TFF	thin film filter
TIA	transimpedance amplifier
TOSAs	transmitter optical sub-assemblies
Tx	transmitters
VCSEL	vertical-cavity surface-emitting laser
VNA	vector network analyzer
VNI	visual networking index
VOA	variable optical attenuator
WDM	Wavelength-division multiplexing
ZB	Zettabytes

Chapter 1

The Role of Optics in Computing

Neural networks and optics were independent entities, until the idea of an optical neural network came up in 1985 [1]. This chapter will cover the background and motivation behind their union. The discussion will start with the challenge inherent to the explosion in data generation/processing and their effects on the global energy consumption. On top of that, conventional digital electronic processor has met challenges, which trigger a demand for new computational architecture. Therefore, the discussion continues with a comparison between different processing machine architectures, including an ingenious brain-inspired neural network. Then, the idea of optics as a promising building block is introduced. Finally, the scope of research will be described to define the objectives of the thesis. Eventually, an outline of the thesis will be presented, including a list of contributions to the field through published research work resulting from the thesis.

1.1 The Edge of Information Age

Our modern society is unavoidably connected to internet. Communication with friends through social media, reading electronic news in mobile phone, storing data in cloud services, ordering goods via electronic commerce, and digital streaming services for entertainment and gaming. These are examples how our lives are attached to a large amount of data. Unsurprisingly, the data volume has grown excessively since the birth of the Information Age in the mid-20th century. The internet traffic rose at an exponential pace to 1.1×10^{21} bytes or Zettabytes (ZB) in 2017 and has greatly influenced the world's power consumption [2]. Cisco predicted that the annual global internet traffic would reach 4.8 ZB per year by 2022 [3]. Figure 1.1(a) shows the expected impact in years based on four main categories of information communication technology (ICT) such as networks (including wireless and wired), data centres (storage and computation usage), consumer devices (televisions, mobile phones, computers, etc.) and the productions of the respective appliances for all the aforementioned categories).

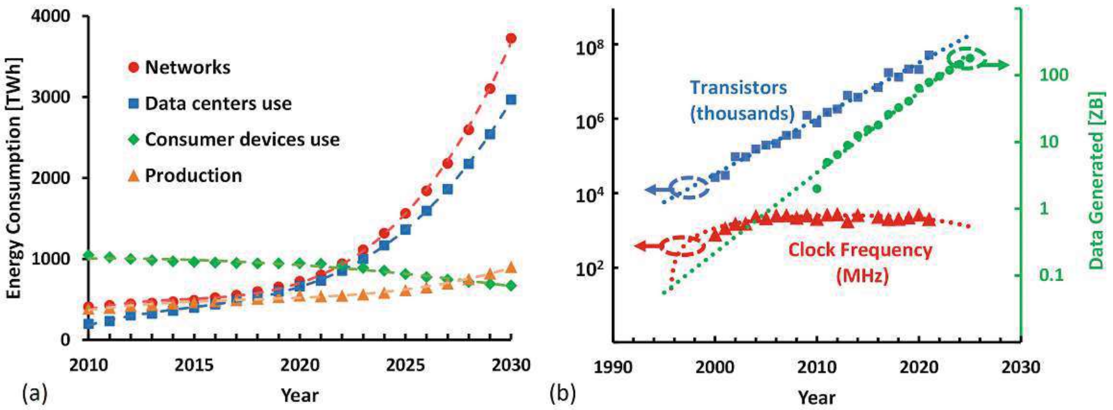

Figure 1.1: (a) Expected impact of information communication technology (ICT) to energy consumption during recent years. Data is sourced from [4]. (b) Comparison of data generated (ZB), numbers of transistors or Moore's law (thousands) and computer clock frequency (MHz) in logarithmic scale through years. Data is sourced from [5]-[6].

To give an idea of the impact, ICT was responsible for 3-4% global human induced carbon dioxide (CO_2) emissions in 2020. This is equivalent to 24 nuclear reactors or the footprint of the commercial aviation industry (900 billion kilograms of CO_2) [2]. Based on Fig. 1.1(a), the use of electricity by consumer devices (◆) is reducing and vastly being taken over by the networks (●) and data centers (■). It can be seen that the power consumption of networks and data centers will keep growing, following an exponential pace. Moreover, the forecast [4] has projected that in the worst case, the total electricity demand due to ICT could lead to a high share of up to 23% in the global energy footprint by 2030 if no measures are taken to mitigate the energy burden of ICT systems.

Apparently, 8% of the ICT portion to the global emission is coming from hyperscale data centres to accommodate an excessive data generation. Hyperscale data centre refers to information factory having an ability to scale the computing infrastructure efficiently up to hundreds of thousands of servers. This trend emerged about a decade ago when companies such as Amazon and Google began to build more data centres. Currently, there are 5620 data centers in the world [7] and the number is expected to increase in the future. Therefore, beyond a doubt, the ICT could share one-fifth of global energy consumption in 2030, as predicted by the forecast. It appears that the term "Data Age" is recently arising as the world is inevitably being controlled by data [8].

Moreover, Moore's law, referring to a doubling in the number of transistors in a computer chip every two years, has governed the evolution of computation technology. Figure 1.1(b) shows a comparison of data being generated, the number of transistors (Moore's law), and the computing efficiency (clock frequency) in logarithmic scale, against the years of technology development [9]. It can be seen that the growth of transistor density (■) follows the same pace as that linked to data generation (●), meaning that Moore's law is still progressing to fulfil the demand of computation. However, slightly before 2010, the computing efficiency began to plateau, due to a saturation in central processing unit (CPU) clock frequency (▲) [10]. Therefore, it is apparent that while the data keeps growing, the current computation efficiency cannot sustain the demand in data processing. Even though the computational machine can be clocked at its maximum frequency, it will induce an excessive power consumption, accompanied by significant heating. This "brick wall" in computation efficiency is believed to culminate in an end of Moore's law [10], leading to a computing gap where a new technology is required to fill in. In the next section, a cutting-edge architecture to overcome this important challenge will be discussed.

1.2 Bio-Inspired Neural Networks and the Role of Artificial Intelligence

The human brain is an important element in human life and yet it is considered a quite complex system in the universe. The human brain hosts about 10^{11} neurons; each of these is connected with $\sim 10^4$ inputs, resulting in a total of 10^{15} synaptic connections [9]. It has extraordinary abilities to think, learn, adapt, and decide. These capabilities are carried out with 10^{13} to 10^{16} operations per second with a power of only 15-25 W [11]. This corresponds to a computational efficiency that is eight orders-of-magnitude higher than human-created digital machines [12]. Due to this astonishing advantage over the classical computers, a bio-inspired artificial neural network (ANN) model has been an integral part of modern technologies concerning artificial intelligence (AI).

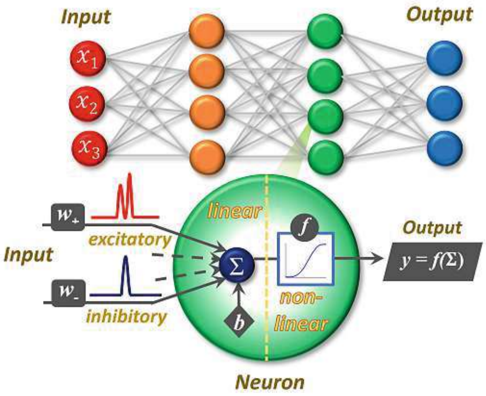

Figure 1.2: Illustration of a neural network.

An example for a conventional ANN architecture is shown in Fig. 1.2, consisting of two fundamental operations. In the first operation, each of neurons x_i ($i = 1, 2, 3, ...$) is weighted with a real number w_i in a range within [-1,1]. The positive (w_+) and negative (w_-) weights give rise to an excitatory and an inhibitory inputs, respectively. The multiplication between neural input and weight is inherently a linear combination, referred to as a multiply accumulate (MAC) operation, resulting in a summed value Σ. This operation is also known as a linear weighted summation. Afterwards, a bias b is added to the Σ, before executing the second operation.

In the second operation, the weighted sum Σ undergoes a thresholding process via a nonlinear activation operation. The activation function is usually bounded and monotonic. If the sum in a neuron exceeds a certain threshold, the neuron is activated, meaning that the sum is forwarded to another node in a subsequent layer of ANN. This is analogous to a signal being transferred to the next neuron in a biological model. On the other hand, if the sum is below the threshold, it is not passed to the next layer, as the neuron's output is zero.

Software-based brain-inspired NNs become popular for solving day-to-day tasks of otherwise computationally hard problems. However, training an AI is costly for most of the complex deep learning models. Based on the trend, training a single deep neural network could consume the CO_2 emissions of five cars over their lifetime [13]. However, with the demonstration of a successful coding algorithm for solving complex problems, tremendous efforts have focused on the development of intelligent machines based on ANNs, aiming at an energy-efficient information processing machine. In the next section, comparison between this "neuromorphic" computing method and other architectures will be discussed.

1.3 Computing Methods

One cannot easily turn away from traditional computing machines developed based on von-Neumann architecture. A continuous development of digital electronic processing machine infrastructure to the level of supercomputer helps in solving complex issues. Furthermore, quantum computing has emerged since few years back and becomes popular, following Google revealed its breakthrough with quantum supremacy [14]. However, neuromorphic computing is becoming intriguing due to its ability mimicking human brain. The characteristics of each processing machine technology are discussed in the following.

1.3.1 Von-Neumann Computing

Conventional digital electronic computers are based on von-Neumann architecture [15]. It consists of a memory that stores both data and instructions, a CPU, inputs, and outputs

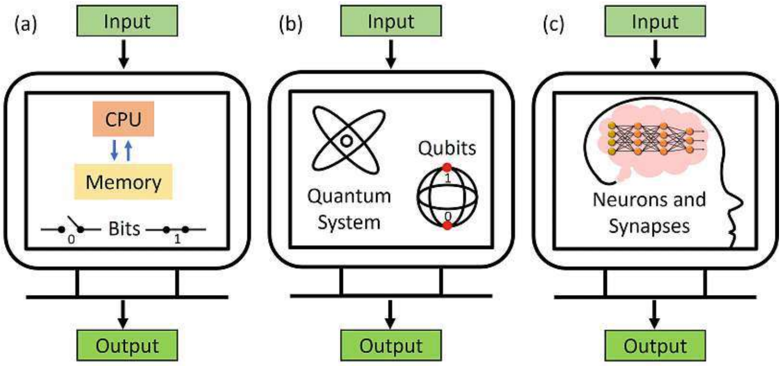

Figure 1.3: Illustration of computer architectures: (a) von-Neumann computing, (b) quantum computing, and (c) neuromorphic computing.

(Fig. 1.3(a)). In this architecture, the communication goes sequentially as data moves between the physically separated processor and memory. Furthermore, von-Neumann computers encode information represented by binary values or bits (0 and 1). However, since instructions are executed one at a time between memory and CPU (point-to-point communication), a latency is unavoidable. Moreover, the data transfer rate (~50 Gb/s) between the processor and memory is much lower than the speed of CPU (~400 Gflop/s) to process data itself. This is known as a von-Neumann bottleneck, as the throughput is the computing limitation. With no improvement in data transfer rates, the high-speed processor spends more time idle, waiting for data to be fetched from memory. Thus, the efficiency of current computer is slowing down, setting up a digital efficiency wall.

1.3.2 Quantum Computing

Quantum computing works on the principle of superposition and entanglement of quantum bits (qubits), as illustrated in Fig. 1.3(b). A qubit state can be in a superposition of 0- and 1-states at the same time. Furthermore, entangled qubits are correlated as they cannot be described independently from each other. These two properties allow quantum computers to attract significant attention because of their capability to solve some computationally-hard problems for classical computers, such as factorization of 1024-bit number using Shor's algorithm [16]. There are several physical systems that can create platforms for a quantum computer, such as superconducting, trapped ion, nitrogen-vacancy (NV) centre, and topological qubits [17]. However, most platforms need temperatures close to absolute zero. Furthermore, the qubits need to be shielded from experimental environments to avoid dissipation and decoherence, as these effects would result in a loss of quantum state population and destruction of state correlations.

1.3.3 Neuromorphic Computing

Neuromorphic computation attempts to mimic the way the human brain works, creating a device that can learn, retain information, and make logical deductions. In a neuromorphic-based computer, both processing and memory are governed by neurons and synapses, as depicted in Fig. 1.3(c). It is different from von-Neumann architecture, since the neurons and synapses can operate simultaneously (parallel processing) [18]. Furthermore, neuromorphic computer receives spikes as input, where the occurring time, their magnitude, and shape are used to encode information. In contrast to quantum computing, the neuromorphic computing architecture is not excessively demanding with respect to its technical implementation, as it can easily work under room temperature conditions. This increases its practicality and potential to be integrated in a wide range of platforms. Furthermore, it requires a much lower power consumption. Thanks to its parallelism principle, the data processing will in general not show a large latency.

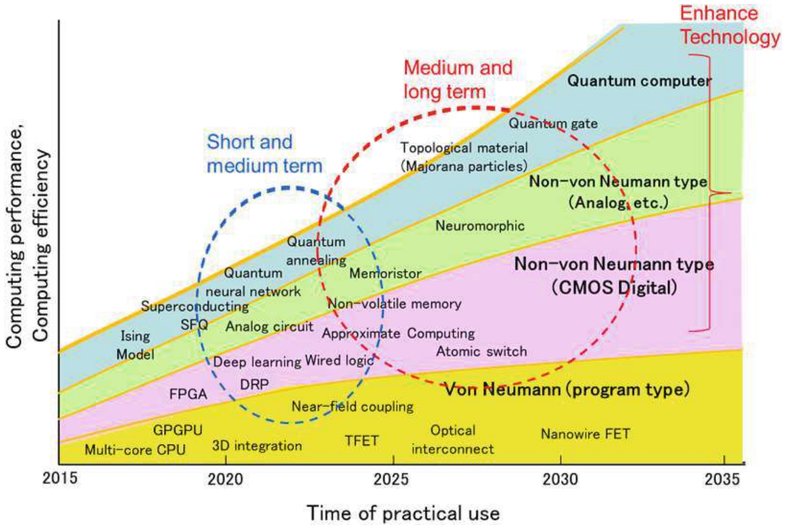

Figure 1.4: Prediction of research and development (R&D) of computing technology over years. Figure is reprinted from [19].

1.3.4 Comparison among the Compute Architectures

In line with Section 1.1, the von-Neumann architecture is bounded by latency obstacle. Quantum and neuromorphic processing machines become favourable candidates as a next-generation computation technology. Figure 1.4 shows the research and development R&D landscape concerning computing technology with respect to a tentative timeline (in year), based on prediction [19]. It can be seen that quantum computers are anticipated to be an efficient technology on a long-term goal (over 10 years). However, in reality, it is difficult to achieve a system free of decoherence, which is the main challenge of quantum computers. Therefore, it will take a long journey to acquire "real" and universal quantum computers. In fact, they will work in a hybrid architecture with classical computing in most cases, or more likely to be specialized machines applied to specific types of problems they are uniquely suited to solve. On the other hand, it is noticeable from Fig. 1.4 that neuromorphic computing is undoubtedly having a promising computing performance that is apparently achievable as a medium-term goal (over 5 years). It is worth to mention that the R&D in neuromorphic computing has been actively developed over years and its state-of-the-art will be discussed in the following section.

1.4 State-of-the-Art in Neuromorphic Computing: Electronics vs. Photonics

The origin of neuromorphic engineering started in the 1980s, when Carver Mead published his research on an analogy between properties of transistors and a neural network [20]. Since then, neuromorphic computing implementations based on digital circuits have grown due to a reliable complementary metal-oxide semiconductor (CMOS) technology. However, some challenges cannot be addressed through this approach, as will be discussed in the following. Note that optical computing has turned into an uprising star, when initial investigations on photonics neural networks were conducted [21]-[22]. The strong points of photonics as neuromorphic computing platform will be reported in this section.

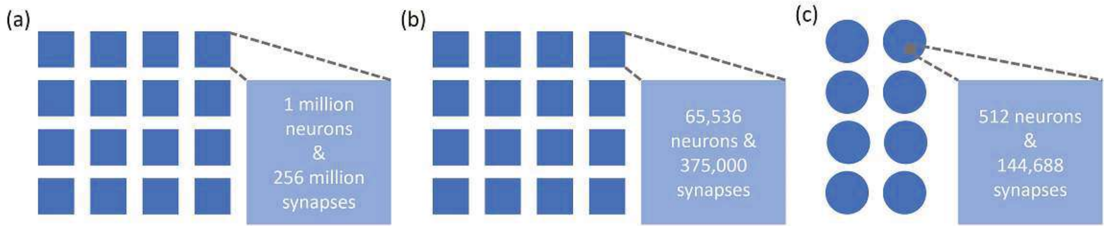

Figure 1.5: Illustration of state-of-the-art of neuromorphic electronics from (a) TrueNorth from IBM, (b) Neurogrid from Stanford University, and (c) HiCANN dies as building block of BrainScalesS wafer from University of Heidelberg.

1.4.1 Neuromorphic Electronics

There are a number of large-scale neuromorphic electronic systems that have been demonstrated over the past 14 years. TrueNorth from IBM [23] features a 28 nm CMOS chip that confines 5.4 million transistors and incorporates 256 neurons, as illustrated in Fig. 1.5(a). Each neuron is connected to 256 synaptic inputs giving out 4,096 neurosynaptic cores. The TrueNorth chip consumes only 20 mW/cm^2. Such low-power neuromorphic chip has been demonstrated as sensor, aiming a real-time cognitive application [24].

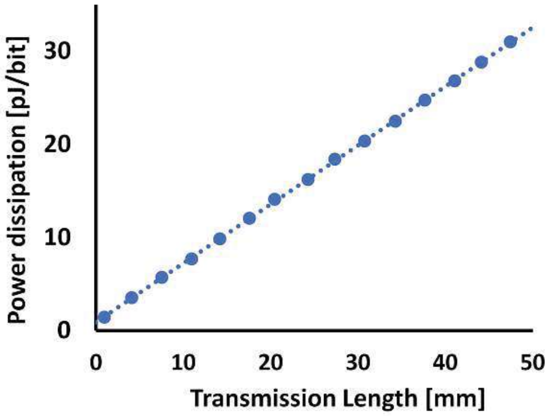

Figure 1.6: Theoretical power dissipation (pJ/bit) of electronic interconnects over distance (mm). Data is taken from [25].

The second example is Neurogrid from Stanford University, with a hardware that comprises 16 chips where each chip has 65,536 neurons with 375,000 synapses [26] (Fig. 1.5(b)). The Neurogrid demonstrator has been tested to simulate the biological brain which has potential for autonomous robots in the future.

Another example is coming from the University of Heidelberg, where they have developed the BrainScaleS neuromorphic system. It uses the wafer-scale integration to have a large number of interconnected neurons. A BrainScaleS wafer has 48 reticles, each of them holding eight High-Count Analogue Neural Network (HiCANN) dies [27]. Each die consists of 512 neurons, including 144,688 synapses, as shown in Fig. 1.5(c). The BrainScaleS has been demonstrated to run a complex modelling task with 10^4 faster than on a classical computer, which is promising for emulating long-time learning application[28].

By integrating a large number of transistor, these examples have shown the possibility to dramatically shrink neuromorphic hardware to chip-scaled size. However, the signal processing on this platform relies on slow-timescale operation to accomplish dense interconnection. On the contrary, a large number of interconnections is inherently required by a neuromorphic processor.

Figure 1.6 shows the theoretical energy dissipation of electronic interconnects over signal transmission length [25]. It can be seen that as the distance increases, power consumption grows linearly. However, to realize high-bandwidth communications, a lower energy consumption can only be realized at a smaller distance, which confines signal transmission to the adjacent chip neighborhood [29]. Beyond this range, conventional electronic links operating at high frequency face difficulties in maintaining a sub pJ/bit efficiency due to various effects that lead to higher losses and excitation of higher-order parasitic modes. This corresponds to the discussion in Sections 1.1 and 1.3.1 about the digital efficiency wall. Therefore, the development of neuromorphic computing cannot rely solely on an electronics platform in the long term.

1.4.2 Neuromorphic Photonics

Photonics is a suitable building block to turn the notion of neuromorphic computing into reality [30]-[31]. It drives ANN processing, thanks to its inherent abilities to efficiently multiplex signals and further due to its low power consumption in view of large electro-optic bandwidths [32]. Optical neural network (ONN) has been conceptually proven to offer low latency and fast signal processing in the GHz range for a wide range of applications [33]-[34]. Neuromorphic photonics represents a broad domain of applications where low-latency, highly parallel, and robust systems become a necessity. The state-of-the-art of ONN demonstrations to realize the linear and nonlinear operations in the photonic domain will be discussed in the following.

1.4.2.1 Linear Operation in an Optical Neural Network

Linearity is at the heart of optics; therefore, such weighted summation has been demonstrated for a plethora of approaches and can be divided into three categories [35]:

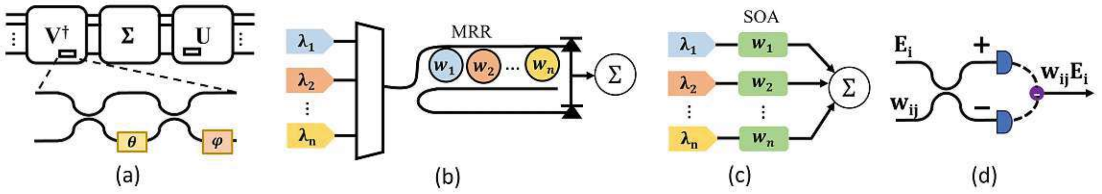

Figure 1.7: State-of-the-art of linear weighting summation function in optical neural network: (a) reconfigurable circuit with Mach-Zehnder interferometer (MZI), (b) broadcast-and-weight neural network, (c) cross-connect-based neural network, and (d) photoelectric-multiplier-based summation.

1. Space-division multiplexing (SDM):

   In this architecture, the neural nodes, including the weighted synapses, are mapped onto the spatial division. The weighting process is done by manipulating the optical fields of the neural data signal. The sum operation is achieved through constructive or destructive interference. Some examples are described in the following.

- MZI-based ONN:

  One example of the SDM-ONN is a programmable integrated photonic based on Mach-Zehnder interferometer (MZI) mesh [36]-[37], as illustrated in Fig. 1.7(a). With a mature photonic integration technology, especially in the field of silicon photonics, the photonic integrated circuit (PIC) offers a high flexibility to scale a complex network. However, the network is conceived for operation at a single wavelength and the control of the phase is usually limited to kHz to MHz range.

- ONN based on grating coupler:

  This work mainly proposes the linear weighted sum based on array of grating couplers integrated on a PIC [38]. An array of 5×6 pixel of grating couplers are distributed as input nodes and used to capture the target images. The weighting process is performed by tuning the input voltage of the PIN attenuators, which are used to alter the optical power of each input. However, the speed of the ONN is still limited by the bandwidth of modulator, amplifier, and photodiode elements on the chip.

2. Wavelength-division multiplexing (WDM):

   Here, the ONN can be demonstrated by employing wavelength–sensitive elements such as micro-ring resonator (MRR) [39] and semiconductor optical amplifier (SOA). The examples are described in the following.

   - Broadcast-and-weight networks:

     Multi-wavelength operation has been applied by dedicating different wavelengths to different input signals while using the wavelength-specific transfer function of MRRs to yield a compact weight bank [40], as shown in Fig. 1.7(b). Apart from its compatibility with WDM, however, the summation of positive and negative weights necessitates the use of both, the drop and the pass-through ports of the resonator, in combination with a balanced photodetector. Therefore, the complexity of this approach is impractical for a large-scale integration.

   - SOA cross-connect:

     SOAs can be exploited in combination with the arrayed waveguide gratings (AWGs). The SOA is used as a pre-amplifier to set the weight (Fig. 1.7(c)), while the AWGs are used to filter out the noise after the cascaded SOAs. However, the SOA-based weight elements are limited by the saturable and noisy non-linear transfer function, accompanied by increased levels of crosstalk due to cascading multiple stages of SOAs [41].

3. Time-division multiplexing (TDM):

   Another approach of ONN is by sequentially map the input data into the time domain. There are some works presented in this area described in the following.

   - ONN for a large-scale fan-in/-out:

     The work [42] is done by encoding both the input vector matrix E_i and the weight matrix w_{ij} onto optical signals by using free-space optical components (Fig. 1.7(d)). The MAC of synaptic connections is realized by the quantum photoelectric multiplication process in the homodyne detectors. This allows a large-scale ONN implementation, where the neural network can be trained and reprogrammed at a high speed. However, such an ANN implementation experiences scaling problems arising from phase instability and aberration of free-space optics.

   - RC-based ONN:

     In the reservoir computing (RC), there are input, reservoir and output layers. The reservoir is built from the recurrent neural network (RNN) model and known as an efficient algorithm. Inputs are mapped into the reservoir with a high-dimension state. The weight connections between the input and the reservoir are randomly generated and remain the same during training. However, the synaptic weights in the output layer are trained. Combining the memory and spatio-temporal processing of RNNs, RC offers the ability to process temporal information. The recent optical RC hardware have been developed based on time delay feedback loops, using vertical-cavity surface-emitting laser (VCSEL) [43] and a silicon MRR [44]. The advantage of this architecture is that the dimensions that facilitates classifications can be increased allowing parallel tasks processing, hence higher speed. However, it is commonly used for a single wavelength only.

1.4.2.2 Nonlinear Activation Function in an Optical Neural Network

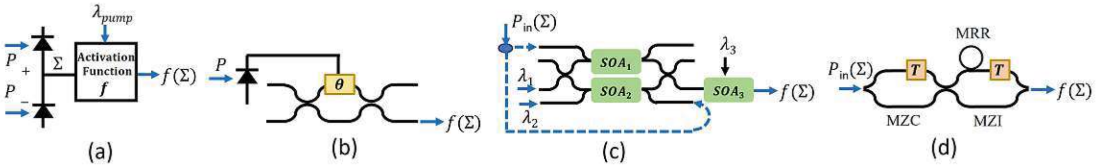

Figure 1.8: State-of-the-art of nonlinear activation functions in optical neural networks: (a) electro-optical modulator, (b) MZI-based electro-optical activation function, (c) SOA-based MZI, and (d) MRR-assisted MZI.

The activation element in a NN plays an important role of discriminating data using a decision boundary, involving functions such as the rectified linear unit (ReLU), the sigmoid or hyperbolic tangent functions – each of these elements can be characterized through different advantages, rendering them as well-suited for a specific set of applications [45]. However, the nonlinear operation of either of those functions is rather challenging to be accomplished in the realm of photonics.

In the past few years, research efforts have demonstrated the feasibility to practically realize optoelectronic or purely optical representations of activation functions [46], as described below:

1. Electro-optical activation functions

    - EAM: One proposal relies on the use of electroabsorption modulator (EAM) [47], as the transfer function of the EAM mimics the sigmoid function (Fig. 1.8(a)) [48]-[49]. The weighted sum (Σ) is converted to a voltage with a photodiode and the voltage is amplified by a transimpedance amplifier (TIA) to drive the EAM. A similar demonstration has been also done by replacing EAM with MRR modulator [50]. By applying different bias conditions, the photodetector-MRR modulator system exhibits a variety of nonlinear activation functions such as sigmoid, ReLU, and quadratic functions.
    - MZI: Furthermore, another study has shown that a MZI-based electro-optical activation function can also introduce nonlinearity function, by modulating the phase and amplitude of the original optical signal (Fig. 1.8(b)), which reveals ReLU and clipped functions [51], [52]. Specifically, in this case, the proposed idea is introduced in the integrated circuit which opens the possibility to scale the complexity of NN.

      A similar proposal can be obtained by different combinations of MRR and MZI on-chip [53], or its combination with optoelectronic detectors [54]. The applied phase shift during modulation is the parameter that modifies the profile of the activation function, including the ReLU-like function.

   However, since all aforementioned nonlinear functions are implemented with a combination of a photodiode and a modulator, these schemes require preceding opto-electronic conversion circuits, which are not energy-efficient.

2. All-optical activation functions

    - SOA-based MZI: An alternative all-optical approach relies on an MZI-like integrated SOA configuration (Fig. 1.8(c)), which is operated in the deeply saturated differentially biased regime, followed by another SOA that acts as a cross-gain modulation-based wavelength converter [55]. However, the device occupies a relatively large footprint, which hinders the scaling for larger scale NN integration.

Similar approaches have been shown by implementing SOA-based wavelength converters in which the synaptic weight is governed by the variation of SOA gain and also realizing a sigmoid function [56], or by employing a self-induced polarization rotation effect in a single SOA [57]. By tuning the initial state of polarization of input signal, the arbitrary activation functions such as ReLU or Softplus can be achieved.

- MRR-assisted MZI: It is possible to employ a plasma effect (Fig. 1.8(d)) by changing the free- carrier concentration in a waveguide to vary the coupling ratio of a Mach-Zehnder coupler to eventually detune the spectral allocation between the MRR resonance and the input signal [58]. Although this scheme enables arbitrary nonlinear activation functions, it is presently bound to latency and a large energy consumption.
- MRR-based activation functions: Another approach of activation function has been realized via silicon MRR loaded with phase change materials [59] or in a Germanium/Silicon (Ge/Si) hybrid structure [60], showing different types of ReLU. However, these approaches necessitate very specific fabrication processes.
- Semiconductor lasers: Further studies have shown that an integrated III-V semiconductor membrane laser can be used to generate a ReLU-like function [61]. By increasing the optical input intensity and the bias current exceeding the threshold value, the light-current ($L-I$) characteristic presenting as the optical output, responds linearly to the input light.

  Another example is to use the bistability of an injection-locked Fabry–Perot semiconductor laser [62]. Nonetheless, these schemes still require external modulation, which complicate their overall integration.
- Periodically-poled waveguide: Compelling results have been shown when exploiting second harmonic generation and degenerate optical parametric amplification as resources, using a nonlinear periodically-poled thin-film Lithium-Niobate waveguide [63]. A linear rectification has been achieved by choosing the length of the waveguide and the bias pulse power. However, the use of a rather exotic material and optical pumps might likely make it difficult to scale up this NN concept.

1.4.3 Electronic vs. Photonic NN Approaches

Neuromorphic electronics is the pioneer of the realization of neuromorphic computing due to the ingenious NN. Figure 1.9 shows that neuromorphic electronics has ~10^4

higher computational efficiency than digital electronics. However, dense and long electronic interconnects lead to a limited signal bandwidth, heat dissipation and high energy consumption. Furthermore, the information bandwidth is limited to the kHz-MHz range due to the latency response inherent to electronic circuits. This could only be matched to applications with a slow information flow, such as real-time voice recognition.

To go beyond the computing speed-wall and digital efficiency brick-wall (Fig. 1.9), neuromorphic photonics is considered as an emerging enabler that combines bio-inspired NNs and photonic integration technology. By capitalizing on the high speed, the high bandwidth and the potentially high parallelism of photonics devices, ONNs offer the possibility to break the limitation of conventional microelectronics when executing MAC operations in complex systems with low latency and a fast GHz signal rates.

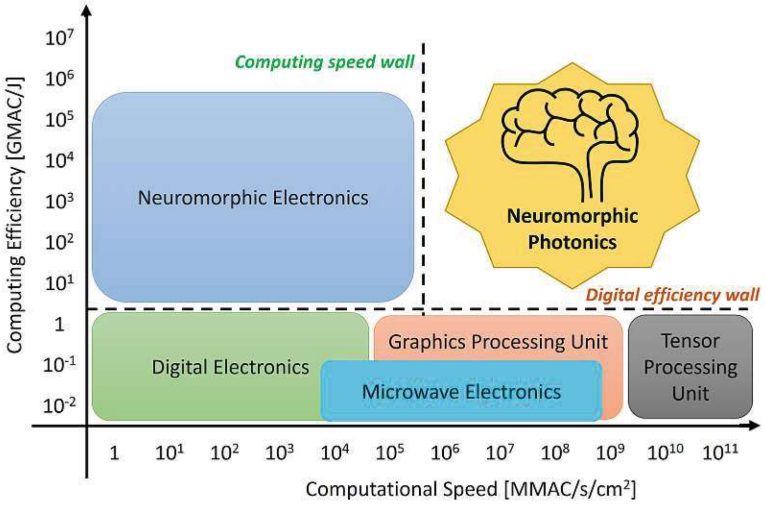

Figure 1.9: Diagram of computational efficiency vs. computational speed for different platforms. Data is taken from [30], [31], [64], [65].

1.5 Challenges for Present-Day Photonic Neural Networks

Until now, the concept of a bio-inspired neural network has been introduced and its marriage with optics is believed to yield a promising future computing architecture. However, there are still two main challenges when it comes to realizing their union.

1. Information bandwidth: The ANN has been initially realized for the field of AI, which has been applied in various applications, such as natural language processing, face recognition, and autonomous vehicles. However, the information bandwidth of AI is limited to the Hz to kHz range, which is not suitable for fast

signal processing applications. In comparison to the electronic-based hardware neural network, the bandwidth spans from kHz to MHz. On the other hand, this bandwidth range is not enough for real-time applications that require fast signal processing up to GHz such as nonlinear optimization, ultrafast control, and quantum tomography. This problem can be tackled by optics because of its vast bandwidth. This enables the use of multiple colors (wavelengths) to send data over the same transmission medium. However, some demonstrations of ONN are still limited to MHz [36], [37]. It is also important to mention that a faster NN response needs to go along with a reasonable energy efficiency. Therefore, a new photonic neural network architecture is necessary to be investigated, finding the middle ground between these demands.

2. Scalability and a flexible synaptic interconnect: High interconnect density is required in neuromorphic computing to perform complex computation. Relying on the electronic-based interconnect is not a feasible solution as its bandwidth is limited and it demands higher power consumption. However, photonics offers a possibility to scale the system to the chip level as fabrication technology becomes more mature. Furthermore, as in a biological neural network, the synaptic interconnect should be compatible with a high fan-in and fan-out, in order to ensure a high degree of connectivity between neurons. In addition, it should also accommodate MAC operation and summation. Thus, a set of analogue operations in the optical domain equivalent to neural network functions need to be investigated to accomplish adequate ONN processing.

1.6 Scope of Research

The focus of research in the thesis is defined as follows, and illustrated in Fig. 1.10

- Flexible synaptic interconnect: The study will focus on realizing colorless synaptic receptor to realize high fan-in of signals, such as required in neural systems. The synaptic interconnect should provide an array of multiplication, where the neurons are being weighted and linearly summed. A combination of suitable and functional photonic components that equate multiplication and addition of high frequency signals is investigated.

- Hybrid opto-electronic and photonic neural network: Besides linear weighted summation, the study will also focus on the realization of the nonlinear activation function in the optical domain. As the next step, the neurons shall be integrated with the activation function to perform the basic NN. Owing to the purpose, the demonstration can leverage on the combination of electronic and optical systems. On top of that, a photonic neuron will be the main focus of the investigation, by

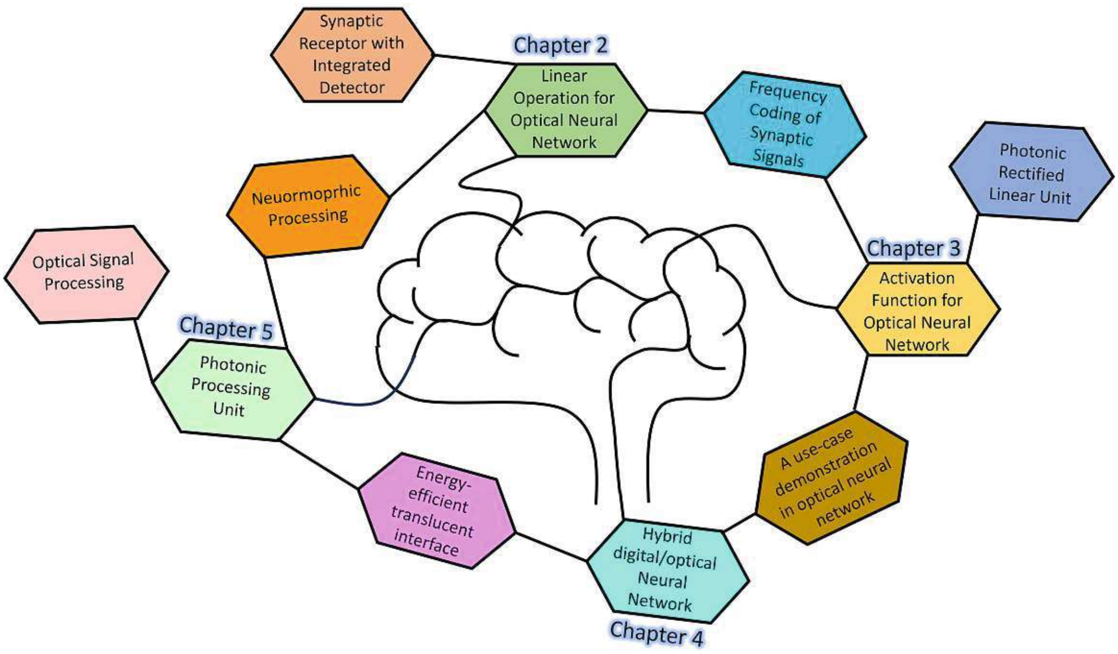

Figure 1.10: Main research topics under investigation in this thesis.

combining the analog weighted summation and the proposed photonic activation function.

- Demonstration of photonic neurons in a use-case problem: To evaluate the performance of the hybrid opto-electronic and all-optical neural network, a benchmark in terms of digital neural network is required. The well-known Iris flower classification problem will be used as a benchmark, given that its NN architecture is convenient for the investigation.

- Integration of multi-purpose photonic processing network: The growth of ONN implementations has triggered the emerge of photonic processing unit research, in contrast to electronic-based graphic processing units (GPUs). Not only the neuromorphic processing can be performed, but also other optical communication network applications can be demonstrated through a flexible layout of the photonic processing unit.

1.7 Outline and Contribution

The outline of this dissertation is presented as following.

Chapter 2 introduces the concept of matrix multiplication in optics via frequency coding. This includes the realization of weighting assignment and the summation. Some potential optical systems that have the equivalent linear operation functions as required for ANN are presented. Finally, the newly discovered colorless synaptic interconnect is demonstrated and its reception performance is reported.

Chapter 3 discusses the optical activation function. Here, the concept focuses on a ReLU and the similar transfer function is offered by an optical bandpass filter in combination with sensitivity clipping, realizing the optical activation operation. Comparison is further made with an electro-optical ReLU function and demonstrated using the use-case of Iris flower classification.

Chapter 4 aims at a hybrid integration of analog linear weighted summation and optical activation function. Here, a new concept of an energy-efficient opto-electronic interconnection between weighted summation and nonlinear activation functions is presented. This contributes to the demonstration of a translucent ONN.

Chapter 5 describes the concept of a flexible photonic processing unit that hosts the ONN. Here, the research emphasizes on the implementation of a photonic processing unit (ΦPU) that includes multiple functions of optical communication networks interconnected in a switch-space architecture such as penalty-free format decoding and neuromorphic functions. In addition, the neuromorphic processing in the switch-space is compared with the one using the wavelength-routed architecture.

The contribution of the thesis is substantiated by the following papers published at scientific conferences and peer-reviewed journals.

[MVS1] M. V. Stephanie, M. Waltl, T. Grasser, and B. Schrenk, “WDM-Conscious Synaptic Receptor Assisted by SOA+EAM”, Opt. Fib. Comm. (OFC) Conf. 2022, San Diego, United States, Mar. 2022, M1G.2.

[MVS2] M. V. Stephanie, F. Honz, N. Vokić, W. Boxleitner, M. Waltl, T. Grasser, and B. Schrenk, “Reception of Frequency-Coded Synapses through Fabry-Perot SOA-REAM Integrating Weighting and Detection Functions”, European Conf. on Opt. Comm. (ECOC) Conf. 2022, Basel, Switzerland, Sep. 2022, Tu5.20.

[MVS3] M. V. Stephanie, F. Honz, N. Vokić, W. Boxleitner, M. Waltl, T. Grasser, and B. Schrenk, “SOA-REAM Assisted Synaptic Receptor for Weighted-Sum Detection of Multiple Inputs (Top-Scored Paper)”, Journal of Lightwave Technology, vol. 41, no. 4, pp. 1258-1264, Feb. 2023

[MVS4] M. V. Stephanie, L. Pham, A. Schindler, M. Waltl, T. Grasser, and B. Schrenk, "All-Optical ReLU as a Photonic Neural Activation Function," 2023 IEEE Photonics Society Summer Tropicals Meeting Series (SUM), Sicily, Italy, Jul. 2023, pp. 1-2, MF4.4.

[MVS5] M. V. Stephanie, L. Pham, A. Schindler, M. Waltl, T. Grasser, and B. Schrenk, "Neural Network with Optical Frequency-Coded ReLU," Opt. Fib. Comm. (OFC) Conf. 2024, San Diego, United States, Mar. 2024, M4C.2.

[MVS6] M. V. Stephanie, L. Pham, A. Schindler, T. Grasser, M. Waltl, and B. Schrenk, "Photonic Neuron with on Frequency - Domain ReLU Activation Function," Journal of Lightwave Technology (accepted).

[MVS7] B. Schrenk and M. V. Stephanie, "ΦPU – A Photonic Processing Unit for Heterogeneous Optical Networks," Journal of Lightwave Technology (accepted).

[MVS8] M. V. Stephanie, L. Pham, T. Grasser, M. Waltl, and B. Schrenk, "Translucent Photonic Frequency-Domain Neuron with Unamplified Concatenation of Neural Sub-Circuits," European Conf. on Opt. Comm. (ECOC) Conf. 2024, Frankfurt, Germany, Sep. 2024, W2A.44.

Chapter 2

Linear Operation for Photonic Neural Network

This chapter will explore the topic of an analog linear weighted summation as one of the building blocks in ONN. The discussion starts with the concept of matrix multiplication of neural inputs and weights in the optical realm. Since processing with optics requires a set of operations equivalent to the NN function, the discussion continues with describing some potential optical systems as building blocks to perform ONN. Furthermore, a wideband monolithic integrated synaptic receptor will be introduced as a promising photonic MAC function. The simulation and the performance of the proposed synaptic receptor are eventually reported.

2.1 Concept of Matrix-Vector Multiplication for the Optically Weighted Summation

As mentioned in Section 1.4.2, there are various ways to perform the optical linear weighted sum of NN since the linear operation is basically at the center of the optical realm. Another scheme that is proposed in this work is employing a frequency-coded signal in the synaptic interconnect, to realize weighting and summation functions. Therefore, the discussion will start with frequency-to-intensity conversion, which plays an important role in weight assignment. Afterwards, various optical transmitters generating frequency-modulated (FM) signals and different optical filters as FM demodulators will be discussed. Finally, the idea of the synaptic receptor with an integrated detector will be introduced.

2.1.1 Frequency-to-Intensity Conversion

Optical fiber technology has revolutionized communication systems by utilizing the optical frequency of light to carry vast amounts of information. This can be done by encoding the digital data into optical signals via modulation. When a laser is modulated, the injected current will vary dynamically and cause the spectral linewidth to be broadened. This signal modulation induces wavelength variation of laser or is known as frequency chirping. Its frequency deviation $\Delta\nu$ is given as follows [66]

$$\Delta\nu(t) = \frac{\alpha}{4\pi}\left(\frac{1}{P(t)}\frac{dP(t)}{dt} + \kappa P(t)\right) \tag{2.1}$$

where $P(t)$ is the optical power of the laser, α is a linewidth enhancement factor and κ is an adiabatic chirp coefficient.

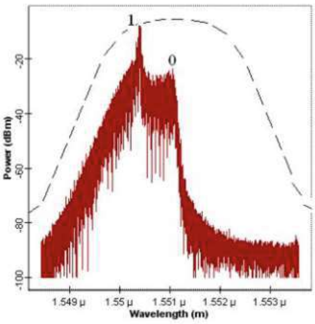

Figure 2.1: Optical spectra of adiabatic chirp-dominated laser (red line) with a Gaussian filter transfer function (black dashed-line). Figure is reprinted from [67].

The first term in Eq. 2.1 is a transient chirp, which is a derivative of the output power resulting in an abrupt change of the refractive index with carrier density. The second part is an adiabatic chirp, which is related to the static power dependence of the cavity refractive index. If the transient chirp is dominant, then the output power will exhibit ringing and overshoot. Conversely, if the adiabatic chirp is dominant, there will be damped oscillations and large frequency deviation [68].

The adiabatic term is the main contributor to frequency modulation of optical signals and causes bit 1 to blueshift relative to bit 0, as illustrated Fig. 2.1. The chirp property of a laser is the foundation to build a frequency-coded signal, as bit 1 and 0 are in two different spectral locations. As an application in ONN, the frequency signal can be encoded such that the mark (1) bit is an active synapse (the spike is emitted) at frequency ν_1, while the space (0) bit represents an inactive synapse at frequency ν_0 (no spike being emitted).

Next, the FM signal can be decoded by employing an optical discriminator to convert the FM signal to an intensity-modulated (IM) signal. FM discriminators have been

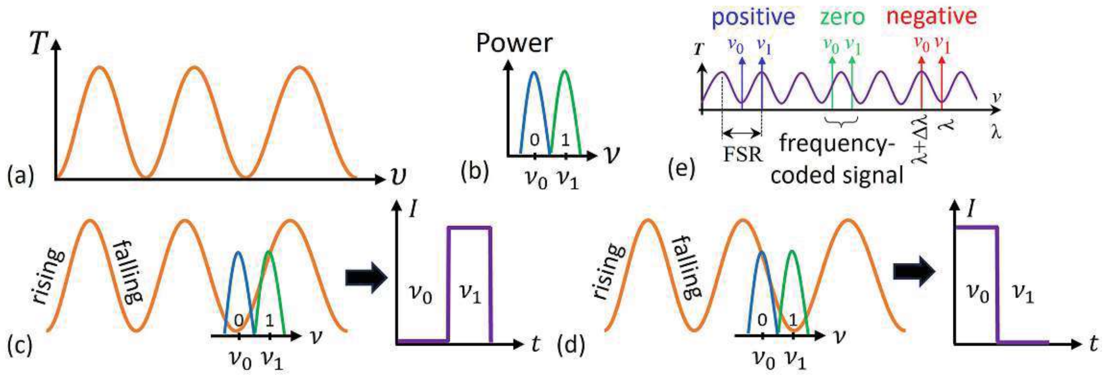

Figure 2.2: (a) A periodic transfer function of an interferometric optical filter. (b) A constant-power FM signal with the lower and upper sidebands are centered at ν_0 and ν_1, respectively. (c) The rising slope filter suppresses the lower sideband (ν_0) of the FM signal and increases the amplitude of the upper sideband (ν_1) after the FM demodulation. (d) On the contrary, the falling slope filter suppresses the upper sideband (ν_1) and increases the amplitude of the lower sideband (ν_0) of the resulting IM signal. (e) The scenario of FM demodulation via the optical filter can be employed for weights assignments to the frequency-coded neural signals [MVS1].

demonstrated theoretically and experimentally by using interferometers [69]-[70]. Additionally, an optical filter can be used as an FM demodulator and typically has a periodic transmission T, as shown in Fig. 2.2(a).

Consider a constant-power FM signal depicted in Fig. 2.2(b), originating from optical carriers centered at ν_0 and ν_1. The lower (ν_0) and upper (ν_1) sidebands represent the space (0) and mark (1) bits, respectively. The FM-to-IM conversion can be performed via the periodic transfer function of the filter, where the ascending or descending slope is centered between the tributaries of the FM signal. Fig. 2.2(c) shows the magnitude response of the ascending slope filter has a positive value. The rising slope is placed between the FM signal and it suppresses the amplitude lower sideband of the FM signal. Therefore, the resulting IM signal amplifies only the upper sideband signal at frequency ν_1. On the other hand, the descending filter slope has a magnitude response of a negative value. Therefore, the falling slope increases the amplitude of the lower sideband of the FM signal and suppresses the resulting IM amplitude of the upper sideband (Fig. 2.2(d)).

This FM-to-IM conversion scheme can be used in ONN to assign the polarity and magnitude neural weight, as shown in Fig. 2.2(e). The filter can be tuned to pass only the mark- or only the space-frequency-coded signal and the IM converted signal is allocated to one sign of the weight, which is eventually sent to the detector. As an example, by suppressing the space (ν_0) component, the resulted IM signal is assigned to positive polarity. On the contrary, by suppressing the mark FM sideband component (ν_1), a negative response is accomplished for the detected signal. Additionally, ν_0 and ν_1 sidebands can be suppressed by the same amount, giving rise to the same transmission during FM demodulation, which can be allocated to a zero magnitude of weight. To apply this scheme effectively, it is necessary to spectrally align the FM signal sidebands to the transmission peak of the periodical transfer function.

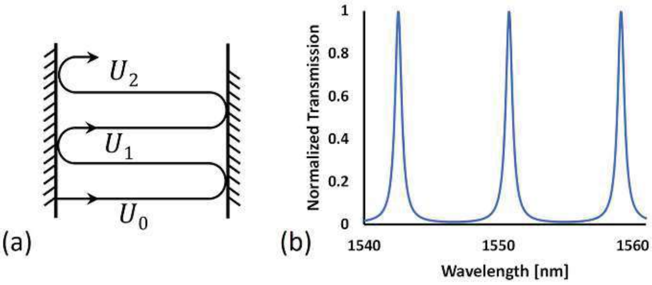

Figure 2.3: (a) Schematic of a Fabry-Perot resonator and (b) its transmission function.

In order to understand the characteristics of a periodic transfer function as in the FM demodulator described above, a mathematical description of the function is provided in the following. For that purpose, a Fabry-Perot etalon (resonator) can be considered. The resonator is built from two-mirror planar as the light wave is reflecting back and forth within the etalon, as illustrated in Fig. 2.3(a). The time-independent wavefunction U propagating inside the etalon can be defined as

$$U(\mathbf{r}) = A\sin(kz) \tag{2.2}$$

where A is an amplitude constant, z is a propagation direction, and $k = 2\pi\nu/c$ is a wavenumber constant [71]. It is relevant to note that this equation satisfies the Helmholtz equation.

$$\nabla^2 U + k^2 U = 0 \tag{2.3}$$

Boundary conditions at the surface of etalon are defined such that $U(z = 0) = 0$ and $U(z = d) = 0$. As a consequence, $k = q\pi/d$ is determined to satisfy the conditions, where q is an integer of the mode number and d is the length of the resonator. In the lossless-resonator condition, the light propagates back and forth infinitely.

However, if the loss is included on the interfaces, there will be some transmitted light. The light still propagates within the resonator, as shown in the Fig. 2.3, and additionally, the reflection constant r is defined here. After one round-trip, the wavefunction will be multiplied with the loss factor h such that $U_1 = hU_0$, where U_0 is the initial amplitude, $h = re^{-j\phi}$ and $\phi = 2kd$ is the phase round trip due to reflection in the mirrors. The superposition of an infinite number of waves becomes

$$U = U_0 + U_1 + U_2 + ... = U_0 + hU_0 + h^2U_0 + ... = U_0(1 + h + h^2 + ...) = \frac{U_0}{(1 - h)} \tag{2.4}$$

Therefore, the intensity in the resonator would be

$$I = |U|^2 = \frac{|U_0|^2}{|1 - re^{-j\phi}|^2} = \frac{I_0}{(1-r)^2 + 4r\sin^2(\phi/2)} \tag{2.5}$$

where $I_0 = |U_0|^2$ is the intensity of the initial wave. In general, the intensity can be defined with a parameter finesse $F = \pi\sqrt{r}/(1-r)$ as

$$I = \frac{I_{\max}}{1 + (\frac{2F}{\pi})^2\sin^2(\phi/2)} \tag{2.6}$$

where $I_{\max} = I_0/(1-r)^2$. F physically describes the relation of the linewidth and the frequency spacing between the peaks. If r is large, then the F is high, meaning that the transfer function has sharp resonances. Figure 2.3(b) displays an example of the linear transfer function of Fabry-Perot etalon for which the parameters of $d = 100$ µm and $r = 0.8$ have been chosen.

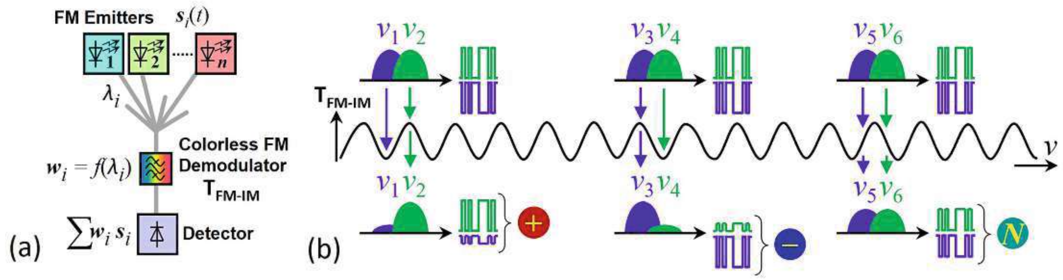

Figure 2.4: (a) Architectural overview of the proposed synaptic interconnect and (b) principal operation of weight assignment through a periodic transfer function for achieving positive and negative signs, including the possibility of setting the magnitude of the weight of having zero magnitude [MVS3].

For the ONN application, the periodic transmission characteristic in Fig. 2.3(b) can facilitate a matrix multiplication between the neural inputs and the weights. The neural inputs (s_i) are translated to FM signals via FM emitters (such as lasers) at different wavelengths λ_i, which are subsequently sent to the FM demodulator, as shown in Fig. 2.4(a). Figure 2.4(b) illustrates the spectra of the FM signal are aligned to the periodic transfer function. Afterwards, the weight assignment (w_i) occurs during the frequency-to-intensity demodulation. The IM signal after demodulation can be assigned to a positive (+) or negative (-) weight, by allocating the spectral position of the mark (ν_2, ν_4, ν_6) and space (ν_1, ν_3, ν_5) bit of the FM signal.

By maximally suppressing one of the tributaries and passing the other, the largest magnitude can be obtained, as illustrated in Fig. 2.4(b) for the positive (ν_1 suppressed, ν_2 passes) and negative (ν_3 passes, ν_4 suppressed) cases. The zero magnitude of weight can be achieved when both FM tributaries experience the same transmission during FM-to-IM conversion, as depicted for ν_5 and ν_6 denoted as N for the result in Fig. 2.4(b). Finally, the received signals at the photodetector after FM demodulation will be straightforwardly summed, eventually yielding the linear weighted addition of neural signals $\Sigma w_i s_i$ (Fig. 2.4(a)).

It is important to mention that the spectral alignment can be performed from the FM emitter side by shifting the wavelength. Practically, the magnitude and polarity of the weight can be set by detuning the periodic transfer function via various parameters, depending on the characteristic of each optical FM demodulator. The next sub-section will discuss some potential FM emitters and FM demodulators as ingredients to perform the linear weighted summation in the optical domain.

2.1.2 Optical FM Transmitters and Optical Resonators

As described in the beginning, the FM signal can be generated by modulating the laser via current injection. In the following, examples of FM transmitters considered in this thesis (directly modulated laser (DML), chirp-managed modulated laser (CML), and distributed Bragg reflector (DBR) laser) will be discussed. It is worth mentioning that the use of these FM emitters will become more significant when the demonstration of optical nonlinear activation function and hybrid digital/optical NN are presented in the subsequent chapters. This chapter is mainly focused on finding a suitable synaptic receptor, performing the optical linear weighted summation. Therefore, various FM-demodulator-assisted optical filters under investigation in this study (delay interferometer (DI), micro-ring resonator (MRR), and optical tunable filter (OTF)) are discussed afterwards.

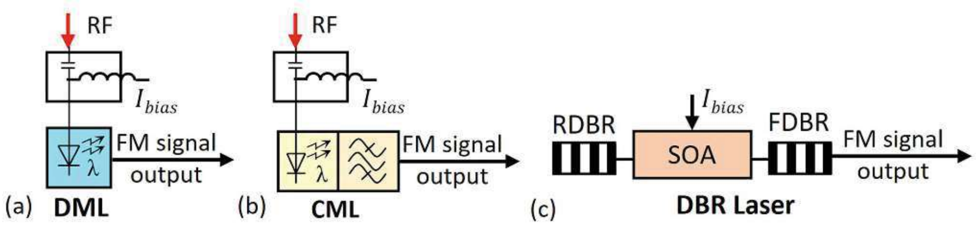

Figure 2.5: Illustration of variety FM emitters: (a) directly modulated laser (DML), (b) chirp-managed directly modulated laser (CML), and (c) distributed Bragg reflector (DBR) laser.

2.1.2.1 Directly Modulated Laser (DML)

As the name itself suggests, an external modulator is no longer required in this laser as the laser diode acts as a light source as well as a modulator. As the light source, the DML is forward-biased by applying electrical bias current I_{bias} and when the I_{bias} exceeds a certain threshold, the laser starts to emit a coherent light. In direct modulation, the input data signal, which is a high radio frequency (RF) signal as the digital data to be transmitted, is superimposed upon the I_{bias}, as depicted in Fig. 2.5(a). This RF signal modulates the I_{bias}, which the variation in the I_{bias} inherently modulates the intensity of the emitted light. The injected current changes according to the input data signal, the output signal varies correspondingly generating an optical signal that carries the encoded data.

The external modulation technique is usually used in optical communication with a longer distance link (>10 km), while the direct modulation is convenient for short-range optical links. The DML is based on the frequency chirping principle, as explained at the beginning of this chapter, where the change of the carrier density via bias current influences the intensity modulation. The DML has a positive chirp characteristic, resulting in a broader laser spectrum that could limit the transmission distance to a maximum of ~20 km for 10 Gb/s and ~2 km for 40 Gb/s signal data rate [68]. However, the issue can be resolved by using a negative dispersion compensation fiber. In spite of that, DML offers a smaller and compact size, low cost, and high output characteristics [72]. It is therefore a well-adopted component in optical communication networks.

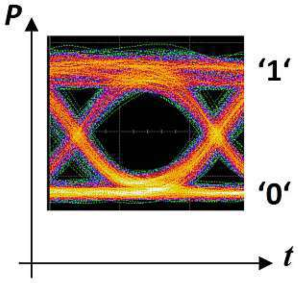

Figure 2.6: Illustration of an eye diagram showing the two power levels of bit 1 and bit 0 of the digital drive signal [MVS3].

The intensity modulation extinction ratio (ER) is an essential performance characteristic for an optical transmitter. The ER describes the ratio of the two optical power levels that represents 1 (higher power level) and 0 (lower power level) bits of a digital drive signal, as illustrated in Fig. 2.6. The ER can be mathematically expressed as

$$\mathrm{ER} = \frac{P_1}{P_0} \tag{2.7}$$

where P_1 and P_0 are the power levels of bit 1 and 0, respectively. In an ideal transmitter, the ER would be an infinite number as P_0 is zero. However, a semiconductor laser is typically biased above the laser threshold ($P_0 > 0$) to sustain lasing conditions and thus a high electro-optic bandwidth. Therefore ER would be a finite number (Fig. 2.6). The quality of the ER can be graphically represented by a so-called eye diagram, as shown in Fig. 2.6. If the eye is widely open, such that the 1 bit at the top can be fairly distinguished from the 0 bit at the bottom, then the quality of the signal in terms of modulation index is good. The DML is commonly biased near the threshold to acquire a good ER between 1 and 0 bits, along with a large current swing to modulate the laser.

2.1.2.2 Chirp-Managed Directly Modulated Laser (CML)

Even though the DML has been demonstrated to have a good performance at a lower cost than external modulation techniques, it has some drawbacks such as a limited transmission reach and waveform distortion due to the chirp effect [72]. Such dispersion effect can be tackled through CML, which is a DML incorporated with an optical spectrum reshaper (OSR) filter, as illustrated in Fig. 2.5(b). By aligning the laser wavelength on the transmission edge of the OSR filter, the red-shifted 0 bits are attenuated and the blue-shifted 1 bits are passed, resulting in an increased extinction ratio (ER) >10 dB [73].

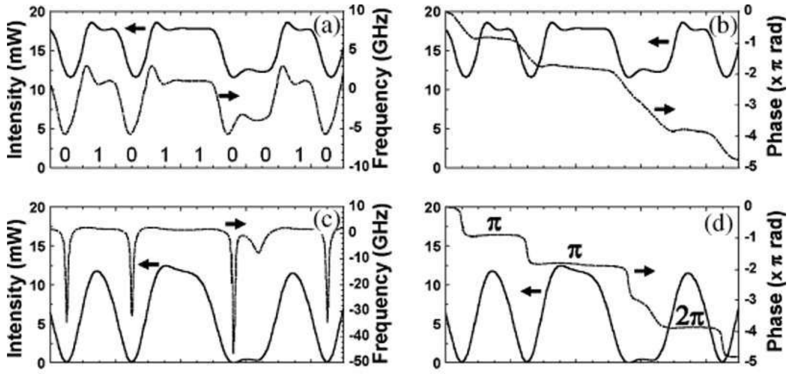

Figure 2.7: Spectrum of CML: (a) Intensity and frequency, (b) intensity and phase before the OSR filter, (c) intensity and frequency, (d) intensity and phase after the OSR filter (taken from [74]).

The OSR filter converts the adiabatic chirp to a flat-top chirp with an abrupt phase shift at the transition of bits, as illustrated in Fig. 2.7 (a) and (c), where the 1 bits are blue-shifted relative to 0 bits. Additionally, there is a continuous π phase shift for each 0 bit sequence (Fig. 2.7 (b) and (d)), meaning that the second 1 bit is π out of phase from the first 1 bit [74]. Thus, it keeps the optical eye diagram open for a long transmission

(>200 km). Furthermore, any remaining transient chirp (amplitude ringing) at high frequency is filtered out, since the bandwidth of the OSR is limited.

Even though the OSR filter increases the dispersion tolerance, it is in fact opposing the goal of generating FM signal since it intrinsically converts the FM to IM signal. Therefore, to generate a FM signal with CML, the OSR filter of CML has to be tuned to a neutral position to avoid frequency-to-intensity conversion in the experimental demonstration.

2.1.2.3 Distributed Bragg Reflector (DBR) Laser

A DBR laser is based on Bragg reflection, where the active medium is embedded between Bragg mirrors. The common three-sections-DBR laser, as illustrated in Fig. 2.5(c), includes a semiconductor optical amplifier (SOA) for optical gain, a front DBR (FDBR) and a rear DBR (RDBR) gratings to form as selective mirrors. A DBR laser offers wide-tunability for WDM applications [75].

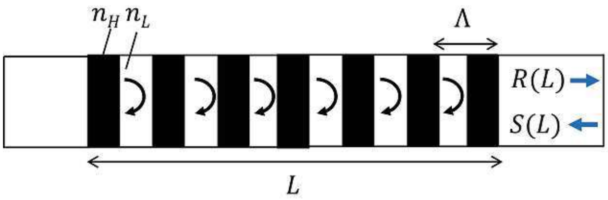

Figure 2.8: Schematic of a DBR grating.

The I_{bias} is injected into SOA to produce a single spatial lasing mode. However, the wavelength shifting can be done by electrically tuning the current injection of both gratings, where the FDBR offers a broad band spectral response with partial reflection and RDBR nearly gives total reflection [76]. It has been demonstrated that a DBR laser can be integrated with an electroabsorption modulator (EAM) into a tunable electroabsorption-modulated laser (EML) [77]-[78]. The SOA and EAM elements will be discussed in detail in the next sub-section of this chapter.

Figure 2.8 shows the scheme of uniform Bragg gratings as the fundamental structure in DBR filter. The Bragg grating consists of a periodic alternating sequence of layers with two different refractive indices. The light reflection occurs between the low (n_{L}) and high (n_{H}) refractive indices. Therefore, in the forward direction of light propagation, multiple reflections occur. The grating period Λ determines the relative phase of all the reflected signals. The Λ is proportional to the Bragg wavelength λ_{B}, defined as

$$\lambda_{\text{B}} = 2n_{\text{eff}}\Lambda \tag{2.8}$$

where n_{eff} is the effective refractive index of the waveguide. Therefore, modulation of the grating can be designed through period Λ and by varying the refractive index of alternating materials so that at the wavelength λ_B, all reflected signals are added constructively (in phase) [79]. At other wavelengths, the reflected light is cancelled out and fully transmitted. The Bragg gratings are governed by the following coupled equations [80]

$$\frac{dR}{dz} = i\delta R(z) + i\kappa S(z) \tag{2.9a}$$

$$\frac{dS}{dz} = -i\delta S(z) - i\kappa^* R(z) \tag{2.9b}$$

where κ and κ^* are coupling coefficients for forward $R(z)$ and backward $S(z)$ travelling modes, respectively, which can be written as

$$R(z) = A(z)e^{i\delta z} \tag{2.10a}$$

$$S(z) = -B(z)e^{-i\delta z} \tag{2.10b}$$

where $A(z)$ and $B(z)$ are the mode amplitude of forward and backward travelling mode, respectively, and δ is a detuning factor

$$\delta = \beta - \beta_B = 2\pi n_{\text{eff}}\left(\frac{1}{\lambda} - \frac{1}{\lambda_B}\right) \tag{2.11}$$

By setting the boundary condition of $R(z = 0) = 1$ and $S(z = L) = 0$ for a grating of length L, the reflectivity will then be

$$R = \left[\frac{S(0)}{R(0)}\right]^2 = \frac{\sinh^2(L\sqrt{\kappa^2 - \delta^2})}{\cosh^2(L\sqrt{\kappa^2 - \delta^2}) - \frac{\delta^2}{\kappa^2}} \tag{2.12}$$

and the maximum reflectivity R_{max} of DBR grating can be achieved when

$$R_{\text{max}} = \tanh(\kappa L) \tag{2.13}$$

The DBR lasing occurs at the wavelength of the maximum reflectivity. The wavelength tuning can be done through the injection current of the gain element or the Bragg grating elements. In the case of an InP-based integrated DBR laser, increasing I_{bias} causes a red shift of the λ_B. On the other hand, increasing the injection current to the RDBR gives results in a blue-shifted of λ_B [79].

2.1.2.4 Delay Interferometer (DI)

A DI has been used for a wide-range of applications, such as quantum key distribution [81], optical sensing [82], modulation format conversion [83], and optical logic functions [84]. Figure 2.9(a) shows the general schematic of DI, which is similar to a

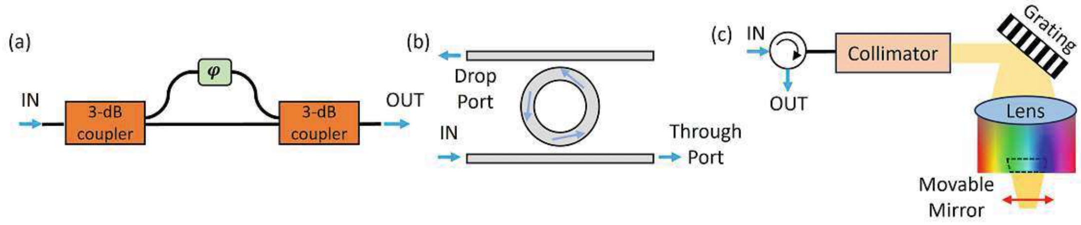

Figure 2.9: Illustration of variety FM demodulators: (a) delay interferometer (DI), (b) micro-ring resonator (MRR), and (c) optical tunable filter (OTF).

MZI. It consists of two 3-dB couplers (50/50 couplers) and they are connected via two optical paths with different lengths. The upper arm has a delay φ which can be altered to tune the output splitting ratio. The periodical transfer function of DI is defined as

$$T_{\mathrm{DI}} = \frac{1}{2} \cdot (1 \mp \cos(2\pi\nu\Delta T)) \tag{2.14}$$

where ν is an optical frequency and ΔT is the time delay in the longer arm [85].

2.1.2.5 Micro-Ring Resonator (MRR)

The MRR is an element realized on a PIC that is fabricated on a waveguide platform. Figure 2.9(b) shows the basic configuration of a MRR, which consists of a ring with radius R and two near waveguides adjacent to it. The ring resonator couples a portion of the light to the adjacent straight waveguide. If the light is sent to the lower waveguide (IN), a portion of the light will propagate along the waveguide and exit via through port. The other portion of light is coupled to the ring and passes to the drop port. The splitting ratio between the through and drop ports depends on the material and geometry of the ring (such as the radius and width of the waveguide) and the optical frequency of the light. The transfer function of MRR from through port can be defined as similar to the Fabry-Perot resonator transmission as following [86]

$$T_{\mathrm{through}}(\nu) = 1 - \frac{1}{1 + \frac{4F^2}{\pi^2}\sin^2\left(\frac{2\pi^2 nR\nu}{c}\right)} \tag{2.15}$$

where n is the refractive index, c is the vacuum speed of light, and ν is the optical frequency. On the contrary, the transmission in the drop port would be $T_{\mathrm{drop}} = 1 - T_{\mathrm{through}}$. The finesse is defined as $F = \frac{\mathrm{FSR}}{\mathrm{FWHM}}$, where the free spectral range (FSR) defines the spacing between two peaks in transmission and the full width at half maximum (FWHM) denotes the width of a line shape at half of its maximum intensity. These parameters are defined as follows [87]

$$\mathrm{FSR} = \frac{\lambda^2}{n_g L} \tag{2.16}$$

$$\mathrm{FWHM} = \frac{(1-ra)\lambda_{\mathrm{res}}^2}{\pi n_{\mathrm{g}} L \sqrt{ra}} \tag{2.17}$$

where λ_{res} is a resonance wavelength, n_g is a group index, r and a are self-coupling and round-trip loss coefficients, respectively.

2.1.2.6 Optical Tunable Filter (OTF)

OTF has been employed for a variety of WDM applications and there are several types depending on the use-cases such as an acousto-optic tunable filter (AOTF), a thin film filter (TFF), a fiber bragg grating filter (FBGF), and a micro electro mechanical systems (MEMS) filter. All of these types are commercially available. However, AOTF, TFF, and FBGF have some limitations such as a narrow tuning range or crosstalk [88], and do not have the freedom of independent control [89]-[90]. On the other hand, an optical MEMS filter is more favourable than others. It has an internal grating and collimating lens, as illustrated in Fig. 2.9(c). The light from the input fiber is collimated and dispersed by the grating which will be focused by lens. A set of wavelengths is reflected from the mirror and couples back to the output fiber. By shifting the position of the moving mirror, the bandwidth and the center wavelength of the filter can be tuned. The optical MEMS filter is attractive because it can be controlled independently and has a negligible chromatic dispersion due to its internal grating [91].

2.1.2.7 Summary

Different types of lasers have been discussed and their function as FM emitters will be explored in the following chapters. The comparisons between DML, CML, and DBR laser are presented in the Table 2.1 [92]. Specifically, the DBR laser here is considered to be integrated with a modulator, as in a tunable EML configuration.

Property	DML	CML	EML
Dispersion limit at 10 Gb/s	5 km	> 200 km	80 km
Maximal launching power	21 dBm	20 dBm	14 dBm [93]
Chirp	Positive transient	Reduced transient & reshaped adiabatic	Negative [94] transient
ER	1-2 dB	10-12 dB	> 20 dB [77]
Cost	Low	Low	High

Table 2.1: Comparison performance between DML, CML and DBR laser with an integrated modulator EAM (EML).

Furthermore, various optical filters have been presented with similar transmission characteristics, even though they are based on different schemes. The comparison between them is presented in Table 2.2. Particularly in this study, the DI characteristics refer to DI from Kylia (MINT product line) with a FSR of 10 GHz [95], the MRR implies a silicon-on-insulator (SoI) MRR [39], and the OTF focuses specifically on a commercial product (Santec OTF-350 [96]) used during later experimentation.

Property	DI	MRR	OTF
Size	Medium	Small	Large
Insertion Loss	2 dB	3 dB	< 7 dB
Tuning range	1.5 × FSR	~ 1 × FSR	80 nm
Tuning speed	ps to nm	µm	-
Tuning mechanism	electro-optic	temperature	manually

Table 2.2: Comparison FM demodulator characteristics between DI, MRR, and OTF

It is worth noting that the frequency-to-intensity conversion does not necessarily have to be accomplished through the FM demodulators presented above. It is possible to perform the MAC operation via alternative optical device solutions which offer a periodic transfer function. Nonetheless, in the next sub-section, a potential synaptic interconnect to perform the optical linear weighted summation is introduced and discussed.

2.1.3 Synaptic Receptor with Integrated Detector

A compact synaptic interconnect is desired to perform the synaptic weighting and summation. One of the prospective candidates is based on the monolithic integration of a SOA and a reflective EAM, yielding an integrated device as referred to a SOA-REAM. This device bears the potential to serve as a colorless demodulator and detector. Each component will be discussed here and their integration as a potential synaptic receptor will be explained.

2.1.3.1 Semiconductor Optical Amplifier (SOA)

Optical amplifiers have been an integral part in the development of optical networks. Due to mature semiconductor fabrication technologies, the SOA has been adopted swiftly as an optical switching element [97], as wavelength converter [55], and for optical phase conjugation [98]. SOA has a basic structure similar to a laser diode, as illustrated in Fig. 2.10. The population inversion between a conduction and a valence band can be achieved through an electrical pumping, creating optical gain [99]. However, besides being operated below the lasing threshold, the most significant difference between a SOA and a laser diode lies in the mirror reflectivity R. The SOA is usually coated with an antireflection mirror, ideally resulting in $R = 0$, to prevent internal feedback. However, there is a residual facet reflectivity as small as $R = 10^{-4}$ practically [100]. Therefore, the SOA gain transmission has a ripple characteristic due to loss (Fig. 2.10) and the resonant spacing depends on its cavity length. Furthermore, a SOA can have a large 3-dB bandwidth gain up to 40 nm [100] and the gain spectrum can be tailored to flexibly center it at arbitrary spectral regions. Due to that, it provides a unique advantage to fiber-based amplifiers that are dedicated to a relatively narrow spectral slice.

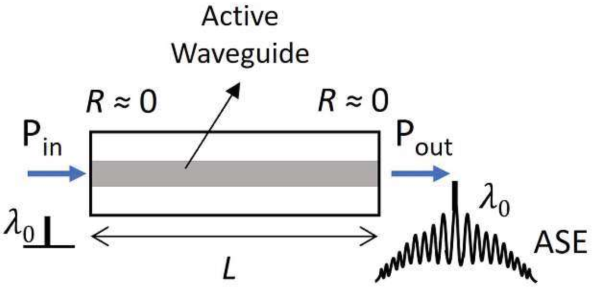

Figure 2.10: Basic structure of an SOA

Theoretically, the gain of an SOA is related to its material gain coefficient per unit length g_m (cm^{-1}), which is proportional to the carrier density (n) and has a parabolic wavelength dependence, defined as [100]

$$g_{\mathrm{m}} = a(n - n_0) - a_2(\lambda - \lambda_{\mathrm{p}})^2 \tag{2.18}$$

where a and a_2 are gain constants, n_0 is a transparency density. λ_{p} is a peak of wavelength, which is proportional to n as following

$$\lambda_{\mathrm{p}} = \lambda_0 + a_3(n - n_0) \tag{2.19}$$

where a_3 is a gain constant. In addition, a net gain per unit length g can be defined in terms of g_m as

$$g = \Gamma g_{\mathrm{m}} - \alpha \tag{2.20}$$

where Γ is an optical confinement factor and α is an effective loss coefficient per unit length. Equation 2.20 is further related to single pass gain G_{s}

$$G_{\mathrm{s}} = \mathrm{e}^{gL} \tag{2.21}$$

where L is an amplifier length. In general, amplifier gain G is related to G_{s} and can be defined as a ratio between output and input signal power

$$G = \frac{(1 - R_1)(1 - R_2)G_{\mathrm{s}}}{(1 + \sqrt{R_1 R_2} G_{\mathrm{s}})^2 + 4\sqrt{R_1 R_2} G_{\mathrm{s}} \sin^2\phi} \tag{2.22}$$

with $\phi = 2\pi n_{\mathrm{eq}} L/\lambda$ and n_{eq} is the equivalent refractive index at zero carrier density. It is noticeable that Eq. 2.22 is very similar to Fabry-Perot etalon transmission function. The ripple along the gain spectrum resembles a comb-like transmission function, as illustrated in Fig. 2.10. Therefore, the gain transmission of the SOA can feature a periodic transfer function provided operation at high gain and with residual reflectivity.

2.1.3.2 Electroabsorption Modulator (EAM)

The EAM plays an important role as an optical transmitter and receiver based on the electroabsorption (EA) effect [47], where the presence of electric field in materials changes the optical absorption coefficient. Therefore, the EA effect offers a dual purpose, namely as an optical intensity modulator with moderate absoption and as an optical photodetector at strong absorption [101].

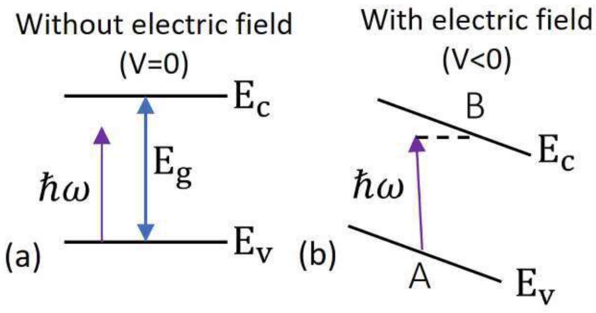

Figure 2.11: Schematic principle of electroabsorption effect in an EAM (a) without and (b) with the presence of electric field.

The EAM waveguide is usually a $p - i - n$ semiconductor stack, consisting of p-type and n-type doped layers, as well as the intrinsic active layer with higher optical refractive index in between to provide a light confinement [102]. The EA effect is based on a Franz–Keldysh effect in the bulk active layer and a quantum-confined stark effect (QCSE) with a multi-quantum well (MQW) in the active region. Figure 2.11(a) illustrates a band diagram of a bulk semiconductor layer without an electric field. When the incoming light is smaller than the energy bandgap ($\hbar\omega < E_g$), an electron in the valence band (E_v) can not make a transition to the conduction band (E_c) since the photon absorption does not provide enough energy. However, when the electric field is applied, the energy bands are tilted, as shown in Fig. 2.11(b). As a consequence, the electron in the E_v at position A can make a transition to E_c at position B because the photon energy $\hbar\omega$ meets the energy difference between the two levels. This scheme is also known as a photo-assisted interband tunneling [102]. The absorption coefficient increases with the applied electric field. The similar principle is also applied to QCSE, where the EA effect occurs in the quantum well [103].

The transfer function τ of the EAM is a function of an exponential of EAM voltage bias as described as [104], [105]

$$\tau = \frac{P(V)}{P_{\max}} = (1 - \epsilon_{\max})e^{-(V/V_{\mathrm{a}})^{a}} + \epsilon_{\min} \tag{2.23}$$

where $P(V)$ is the output power as a function of voltage V, $P_{\max}$ is the output power when zero bias voltage is applied, V_{a} and a are fitting parameters and $\epsilon_{\min}$ is a minimum extinction parameter. Another important parameter in the EAM is a static ER, which is a ratio between the output power at different applied bias and at zero bias ($V = 0$). The static ER is related to the change of absorption coefficient with the applied bias [103]. It is worth mentioning that a static ER up to 9 dB can be achieved by applying a low driving voltage 1 Vpp swing [47], which is beneficial for optical modulation or broadband photodetection. Even a low driving voltage below 400 mVpp has been reported that it could at least reach a 6 dB of ER [106].

Besides its superiority in data transmission, the EAMs are able to translate the signal from the optical to the electrical domain by harnessing the absorption property in a photodetection. As a photodetector, the EAM converts the incoming optical signal that passes through the modulator, providing that it is reverse-biased at absorption. The incident photons are absorbed and generate electron-hole pairs within EAM. The applied electric field accross the EAM causes the electrons to move towards the positively biased electrode while the holes move towards the negatively biased electrode, generating the photocurrent. The magnitude of the photocurrent is proportional to the intensity of the incident light, allowing the EAM to detect and quantify the optical signal.

The photodetection reception ρ has a dependence on a voltage bias V_{EAM} as follows

$$\rho(V_{\mathrm{EAM}}) = 1 - \tau(V_{\mathrm{EAM}}) \tag{2.24}$$

where τ has been defined in Eq. 2.23. It has been showed that a high photocurrent of ~20 mA can be achieved with a low bias of −1.37 V, therefore EAM has a potential as a low-biased photodetector [47]. Additionally, recent work has demonstrated the wide-bandwidth and high responsivity of EAM as a photodetector in the C + L band wavelengths [107].

2.1.3.3 Monolithic Integration of SOA-REAM and its Role in an Optical Neural Network

As described above, the SOA can provide a large spectral bandwidth for its optical gain and the EAM serves as a fast detector when reverse biased. Therefore, the integration of both is promising as an FM demodulator and detector as a so-called SOA-REAM (Fig. 2.12(a)) [108].

The configuration of the SOA-REAM synaptic receptor is shown in Fig. 2.12(b). The front facet of the SOA is partially reflective, while the end facet of the REAM is

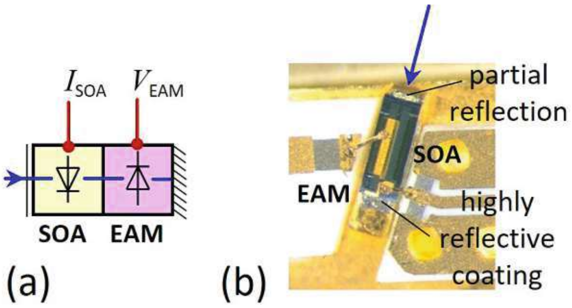

Figure 2.12: (a) Schematic of SOA-REAM and (b) the monolithic integrated SOA-REAM which intentionally features a residual reflection at its input facet [MVS1].

highly reflective. Therefore, round-trip reflections from both facets form a gain ripple, as illustrated in Fig. 2.10. This comb-like gain transmission can be employed for the synaptic weight assignment through the FM-to-IM conversion described in section 2.1.1 (Fig. 2.4(b)). Furthermore, its periodicity can be utilized for weighting multiple input signals simultaneously. In addition, the EAM will directly receive the weighted addition of neural signals. Therefore, the optical linear weighted summation can be performed via the monolithic integrated SOA-REAM.

The wavelength-selective gain spectrum of SOA-SOAM can be expressed in the following as a function the optical frequency of the input signal ν

$$T_{\mathrm{rx}} = \frac{1}{1-R}\left(\frac{(1-R)^2 G_{\mathrm{soa}} a_{\mathrm{eam}}}{(1+RG_{\mathrm{soa}}a_{\mathrm{eam}})^2 + 4RG_{\mathrm{soa}}a_{\mathrm{eam}}\sin^2\left(\frac{2\pi\nu}{\mathrm{FSR}}\right)}\right) \tag{2.25}$$

where R is the front facet reflectivity, G_{soa} and a_{eam} are the SOA gain and the EAM loss, respectively.

The demodulation of frequency-coded synapses via the SOA-REAM fits the architecture illustrated in Fig. 2.4(a). The SOA represents the FM demodulator, providing a weight setting w_i for a given wavelength channel λ_i. This setting depends on the spectral alignment of λ_i relative to the comb-like transfer function T_{rx}. The alignment is governed by tuning either the comb or the wavelength of the input signal s_i. Because the SOA gain ripple enables operation across a broad optical bandwidth, accommodating multiple wavelength channels, the photodetector receives a weighted sum of all synaptic emitters. This is possible through the wide optical bandwidth of the EAM photodiode. Consequently, the SOA-REAM cavity serves as a colorless demodulator, where “colorless” denotes the ability to process multiple channels that are spectrally using the same transfer function T_{rx}, due to its periodic transmission property. It is important to mention that the “colorless” does not refer to the wavelength-agnostic behaviour in terms of transmission, but rather to its capability to process multiple wavelength channels.

The following section compares and discusses the transfer functions of the optical filters and SOA-REAM in more detail.

2.2 Simulation of Some Potential FM Demodulators

The simulations are performed by using MATLAB. This sub-section focuses on the transmission of DI, MRR, and SOA based on their mathematical representation described in the previous sub-section. Afterwards, the comparison between the FM demodulators based on the numerical simulation is presented.

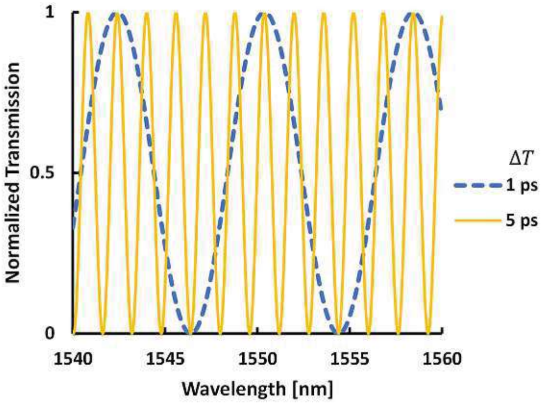

Figure 2.13: Transmission of DI for time delay ΔT of 1 ps (dashed line) and 5 ps (line).

The transmission of the DI depends on the time delay ΔT, as denoted by Eq. 2.14. Figure 2.13 shows the transfer function of DI for different ΔT. If the delay time gets smaller, the spacing between transmission peaks becomes larger, ΔT of 1 ps (dashed line) and 5 ps (line) are inversely proportional to FSR (1 THz and 0.2 THz, respectively).

For integrated waveguide MRR devices, the FSR of the transmission depends on the geometry and the material of the ring resonator. Figure 2.14 presents the linear transmission of MRR for radius R of 10 µm (dashed line) and 30 µm (line), based on Eq. 2.15. A SoI ring resonator with parameters of $n_g = 4.3$, $r = 1$, and $a = 0.9$ is considered here. It can be noticed that the smaller the radius R, the wider the spacing in the transfer function.

In the case of SOA, the transmission characteristic is shown in Fig. 2.15. The simulation is based on Eq. 2.22 and using parameters in Table 2.3. The parameters are taken from [100]. It can be seen that the SOA gain has a comb-like characteristic with a peak extinction of 2.3 dB. In this simulation, a SOA length of 400 µm has been considered, resulting in a FSR of 0.99 nm.

Based on the simulations presented above, the DI, MRR, and SOA can be used as FM demodulators, taking advantage of their periodic transfer functions. Particularly,

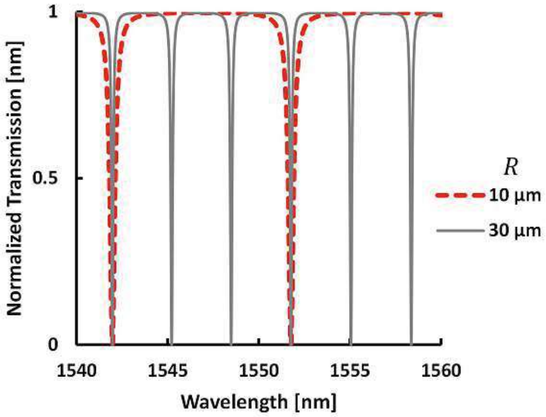

Figure 2.14: Transmission of MRR from through port for ring radius *R* of 10 µm (dashed line) and 30 µm (line).

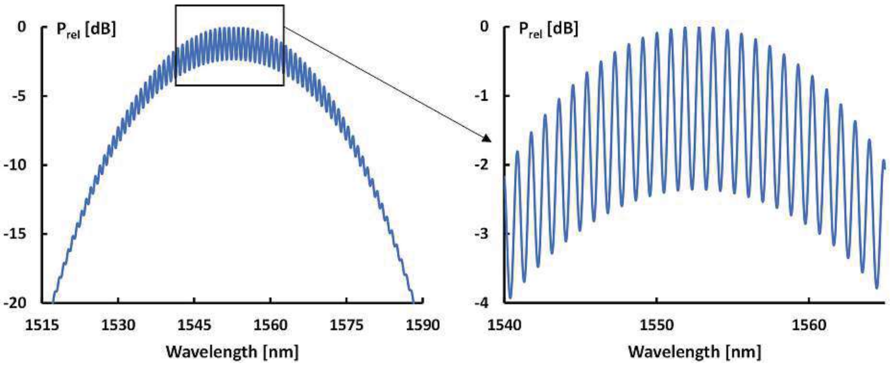

Figure 2.15: Gain spectrum of SOA and its zooming around the ripple

the SOA-REAM is compact and has a small size comparable to the MRR. However, the SOA-REAM has a broader gain spectrum and more ripples between 1515 nm and 1590 nm, as shown in Fig. 2.15. On the contrary, the MRR has about three and seven transmission peaks for the radius *R* of 10 µm (dashed line) and 30 µm (line) as shown in Fig. 2.14, respectively, between the range of 1515 nm and 1590 nm. Additionally, the transmission range of MRR is limited in practice due to the wavelength dependence grating coupler in the input and output of waveguide to extract the response. Due to that reason, the SOA-REAM will be more compatible for WDM application in ONN, considering that multiple inputs at different wavelengths can be simultaneously weighted.

The next section will discuss the elements for evaluating the performance of the SOA-REAM.

symbol	parameter	value
a	gain constant	2.7×10^{-16} cm^2
a_2	gain constant	0.15 cm^{-1} nm^{-2}
a_3	gain constant	2.7×10^{-17} nm cm^{-3}
n_0	transparency density	1.1×10^{18} cm^{-3}
n	carrier density of the gain material	1.2×10^{18} cm^{-3}
Γ	confinement factor	0.3 [109]
α	effective loss coefficient per unit length	50 cm^{-1}
n_{eq}	equivalent refractive index at zero carrier density	3.257
L	amplifier length	400 µm
R_1&R_2	facet reflectivity	0.5
λ_0	center of wavelength	1550 nm

Table 2.3: Parameters considered for simulation of SOA gain transmission.

2.3 Methodology

In order to demonstrate the capability of SOA-REAM as a synaptic receptor, two elements in the experimental setup are required. The first is a synaptic emitter to generate the FM signal and the second is a reception technique to analyze the performance of the synaptic receptor. Each technique will be floored in the following.

2.3.1 FM Signal Generation

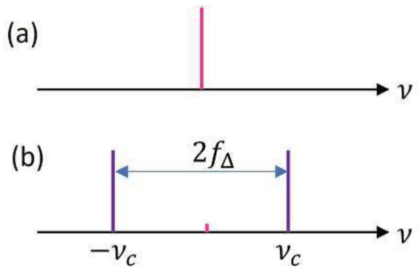

Figure 2.16: (a) The carrier signal spectra and (b) the resulting spectral lines after carrier-suppressed double sideband modulation separated by the modulation bandwidth $2f_\Delta = 2\nu_c$.

To generate the FM signal at two different spectral position, a carrier-suppressed double sideband (CS-DSB) modulation method can be employed, where the frequencies generated by IM are symmetrically positioned above and below the carrier frequency, with the carrier signal ideally being entirely suppressed.

Given the message signal $m(t)$ and carrier signal $c(t) = \cos(2\pi\nu_c t)$ in time domain, the modulation of signal $s(t)$ can be done by multiplying the message and carrier signals, resulting to

$$s(t) = m(t)\cos(2\pi\nu_c t) \tag{2.26}$$

where ν_c is the frequency of the carrier signal. Assuming the message signal $m(t)$ has a Fourier transform $M(\nu)$, the Fourier transform of $s(t)$ is

$$S(\nu) = \frac{1}{2}[M(\nu - \nu_c) + M(\nu + \nu_c)] \tag{2.27}$$

The equation above shows that the resulting signal $S(\nu)$ contains two spectral components, where the first part is in the lower sideband frequency ($-\nu_c$) and the second part is the upper sideband frequency (ν_c). Figure 2.16 illustrates the carrier signal spectra before (a) and after the CS-DSB modulation (b). The carrier signal in the middle is suppressed, while two new lines are formed, which are separated by the modulation bandwidth of $2f_\Delta = 2\nu_c$. Afterwards, both spectral lines can be modulated with a MZM to produce a constant-power FM signal, and the resulting FM signal will play an important role for the weight assignment, as depicted in Fig. 2.2.

2.3.2 Reception Technique

The reception performance of the synaptic receptor can be evaluated in terms of bit error ratio (BER) measurements. The BER measurements are performed to quantify the reception quality of a system by comparing the number of bits received to the total number of transmitted bits, thus counting the number of errors. This can be done by characterizing the reception sensitivity of the receiver, which is the minimum power level that an optical receiver can detect the input signal. This criterion is used to identify the limit of the interfered system. The lower the received signal power, the weaker the signal that the receiver can demodulate to retrieve the information.

Furthermore, imperfections in an optical communication system, such as noise, dispersion, or moderate ER can influence the reception of the receiver and can be compensated by increasing the power. The extra amount of optical power needed to overcome the problem or to obtain the same BER as would be given for an infinite extinction, is defined as reception penalty. The penalty ξ can be defined in terms of ER as following [110]

$$\xi(\mathrm{ER}) = 10\log\left(\frac{1+\frac{1}{\mathrm{ER}}}{1-\frac{1}{\mathrm{ER}}}\right) \tag{2.28}$$

The next section will explore the demonstration of a SOA-REAM as a wideband demodulator and detector. Subsequently, its FM demodulation performance will be compared to the integrated MRR and bulk OTF.

2.4 Experimental Evaluation

This section presents two experiments. The first part demonstrates the SOA-REAM based on a demodulation function of a frequency-coded signal. Here, the characterization of SOA-REAM is presented, and the reception penalty is discussed. The second part evaluates the SOA-REAM as a synaptic receptor with an integrated detector function. In addition, its capability as a synaptic receptor is compared with that of other FM demodulators.

2.4.1 Experimental Setup to Demonstrate SOA-REAM-Based FM Demodulation

Figure 2.17 shows the experimental setup to evaluate the SOA-REAM based demodulation of a frequency-coded signal. Two upstream spike emitters are connected to the downstream synaptic receptor. The spike emitters are based on distributed feedback (DFB) lasers operating at wavelengths of 1576 nm and 1599 nm. Each emitter is connected to a pair of MZM to generate an FM signal through complementary intensity modulation of the two optical seed carriers at λ and $\lambda + \Delta\lambda$, respectively, corresponding to ν_2 and ν_1 in Fig. 2.4(b), yielding an FM signal due to the wavelength switching that is further characterised by a constant optical power.

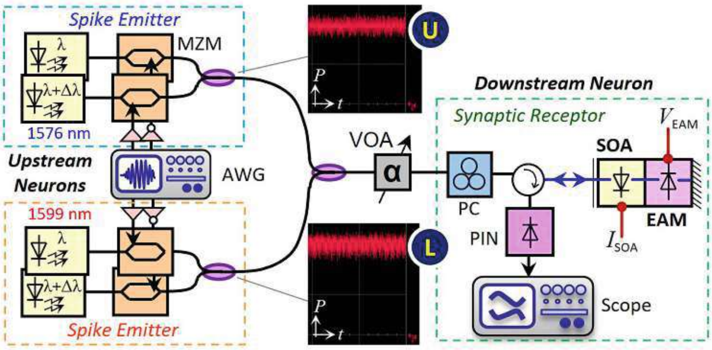

Figure 2.17: Experimental setup to evaluate the demodulation performance of the synaptic receptor and to perform simultaneous weighing and summation of the signals from two different spike emitters [MVS1], [MVS3].

The difference $\Delta\lambda$ between the two optical sources for the same wavelength channel is fine-tuned to the comb function so that $\Delta\lambda = \text{FSR}/2$. The complementary outputs are then combined with a 50/50 combiner and sent to the synaptic receptor. The resulting intensity ER for the generated FM signals (ideally 0 dB) is 0.72 dB for the given transmitter configuration and the constant power output is -1 dBm, as shown in the insets of U and L in Fig. 2.17. It is important to note that the apparently complex MZM-centric

FM emitter configuration described here is used instead of simpler DML / CML based emitters as the focus is solely on the performance evaluation of SOA-REAM as the synaptic receptor.

Next, an arbitrary waveform generator (AWG) is used to drive the MZM modulators with two kinds of waveforms. The first set is a spike train with 100 ps width and a duty-cycle of 1/16. The timing parameter of this spike train is arranged to accommodate the time-interleaving of a second spike train, which is spectrally allocated along the demodulating comb function of the SOA-REAM. The second waveform is a 10 Gb/s pseudo-random bit sequence for BER measurements.

The synaptic receptor comprises the SOA-REAM and a polarization controller (PC) compensate the polarization-dependent response of the EAM (Fig. 2.17). In this experiment, a PIN receiver is employed as an external photodetector due to the unfortunate loss of RF response for the EAM. The demodulated synapses are then digitized using a real-time oscilloscope. To facilitate BER measurements, a variable optical attenuator (VOA) is introduced to adjust the received optical power (ROP) for the receptor.

2.4.2 Weight Tuning and Simultaneous Summation of Signals

The SOA-REAM had a length of 470 µm, resulting in a FSR of 0.55 nm (Fig. 2.12(b)). The transfer function of the L-band SOA-REAM is depicted in Fig. 2.18 against the wavelength. It has an optical 3-dB bandwidth of 23.75 nm with a peak extinction of 4.73 dB for its demodulating comb function. If the spacing of the receptor is determined by FSR of 0.55 nm, the SOA-REAM can accommodate 43 synapses along its spectral bandwidth.

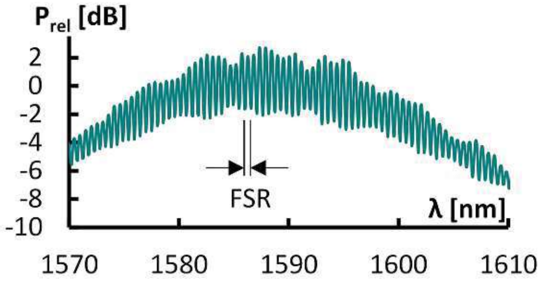

Figure 2.18: Transfer function of SOA-REAM with FSR = 0.55 nm [MVS2].

The relative spectral alignment of the input signal to the periodic transfer function of the SOA-REAM determines the sign and magnitude of the weight of the proposed receptor. The characterization of the spectral tuning of the comb function is reported in Fig. 2.19 by considering several parameters that induce a variation in the optical density of the SOA-REAM cavity. This includes SOA and EAM biases (Fig. 2.19(a)), temperature, and input signal power (Fig. 2.19(b)). The characterization results are shown as the spectral detuning of a local comb resonance around 1585 nm, which is in the middle of the gain spectrum (Fig. 2.18). In the case of SOA bias current variation, the local resonance is blue-shifted until the threshold limit of ~95 mA is reached. For accomplishing a sufficiently high cavity gain, the bias current is linearly red-shifted at a response of 1.34 GHz/mA (● in Fig. 2.19(a)). For the EAM bias, the local resonance has a linear dependence on the bias voltage at a detuning of 9.94 GHz/V (◆ in Fig. 2.19(a)).

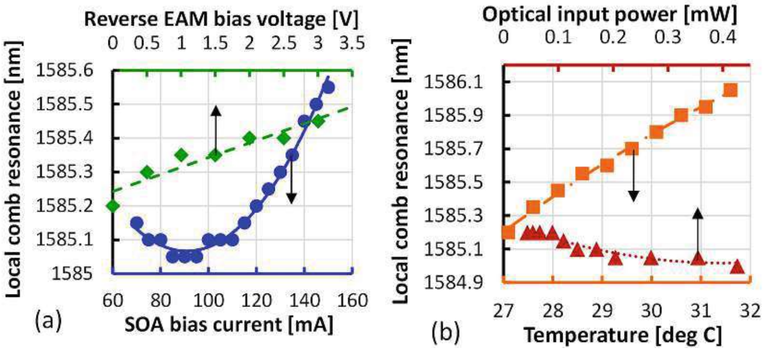

Figure 2.19: Spectral detuning of the SOA-REAM transfer function related to (a) SOA and EAM biases, and (b) temperature and optical input power [MVS1], [MVS3].

In terms of the temperature, the detuning shows a relatively large sensitivity with a response of 22.5 GHz/°C (■ in Fig. 2.19(b)). On the other hand, the optical input power influences the resonance by a spectral shift of −155.6 GHz/mW (▲ in Fig. 2.19(b)), particularly in the range between −13.6 dBm (~0.04 mW) and −3.7 dBm (~0.4 mW). Considering such a small fluctuation in optical input power for synaptic connections with a given optical budget, calibrating the initial receptor can compensate this detuning.

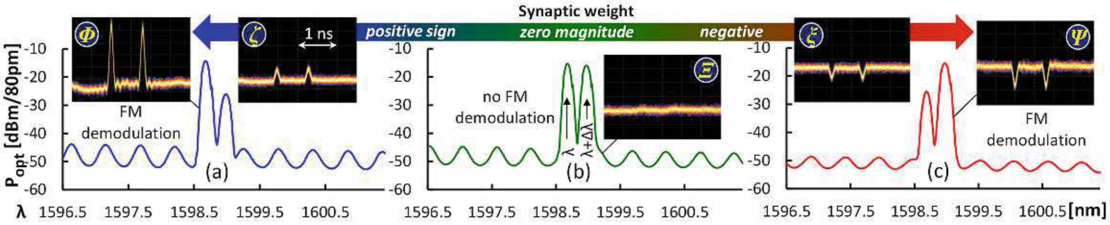

Figure 2.20: Spectra of signal and spike train reception of SOA-REAM with different weighing processes (a) positive sign, (b) zero magnitude, and (c) negative sign [MVS1], [MVS3].

Figure 2.20 shows the optical signal spectra at the output of the SOA-REAM after FM-to-IM conversion for the demodulated FM signal at 1599 nm. The spectra are presented for various weight settings and the corresponding detuning of the comb function. It is worth noting that independent spectral alignment was necessary for at least one of the two spike emitters due to the presence of two emitters in the upstream neuron and one synaptic receptor in the downstream neuron (Fig. 2.17). Therefore, this involves adjustment of the current or temperature of the FM emitter laser.

The insets in Fig. 2.20 show the received spike trains for each weight setting. By suppressing the optical FM sideband at $\lambda + \Delta\lambda$ or at a λ, a positive (Φ) or a negative (Ψ) sign can be obtained for the detected spike train, respectively. Particularly, when no FM demodulation occurs, the received signal has a constant amplitude (Ξ), which is proportional to a zero magnitude of the weight. It shall be pointed out that the magnitude of the weight can be changed by detuning the comb function, as shown for a positive (ζ) and a negative sign (ξ) in Fig. 2.20 for the detected spike trains with reduced amplitude.

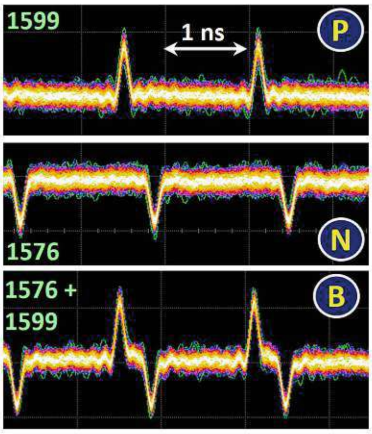

Figure 2.21: Simultaneous reception of weighted spike trains summed by the synaptic receptor (B) where the constituent wavelength channels at 1576 nm and 1599 nm are weighted negatively (N) and positively (P), respectively [MVS1], [MVS3].

In addition, two synaptic emitters have been jointly delivered to the SOA-REAM (Fig. 2.21). In this scenario, the 1599-nm signal has been weighted positively (P) and the signal at 1576 nm is chosen with a negative sign (N). By sending the signals simultaneously from both synaptic emitters at the input of SOA-REAM, the received spike train shows a correct summation (B). This result emphasizes that the gain ripple of the SOA can be used as FM demodulation for performing the weighting function. Furthermore, it is evident that the synaptic receptor can execute larger MAC operations involving multiple synapses at different wavelengths by leveraging the WDM dimension.

2.4.3 Reception Penalty

The reception penalty of SOA-REAM performance is investigated as the next step. Figure 2.22 describes the BER for the two wavelength channels as a function of the ROP at SOA-REAM. For each of the two wavelength channels, the results are presented for a positive (suppressed $\lambda + \Delta\lambda$) and a negative (suppressed λ) weight. A comparison is made with the reception of an IM signal, which is equivalent to the ideal case of an infinite ER for FM-to-IM conversion. The IM reception means only the FM tributary at λ is modulated at the spike emitter.

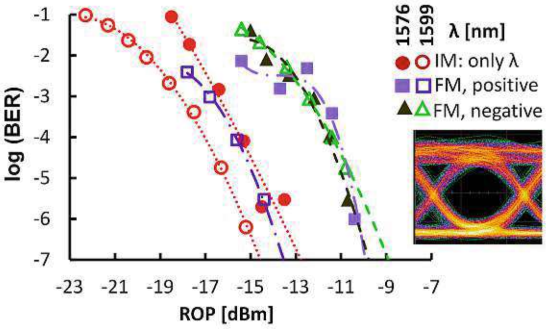

Figure 2.22: BER performance of the SOA-REAM serving as a colorless demodulator for frequency-coded synapses [MVS1], [MVS3].

Despite the moderate FM demodulation extinction of SOA-REAM, as mentioned in the Section 2.4.2, Fig. 2.22 shows that in principle, a low BER can be achieved. In the case of the first channel at 1576 nm (Fig. 2.22), a reception sensitivity of −10.4 dBm is obtained at a BER of 10^{-6} after FM demodulation, with a small deviation of 0.1 dB between the positive and negative weight settings (■, ▲). A reception penalty of 3.2 dB is found with respect to the IM signal (●). This comes from the high intensity ER of the MZM transmitter, as it does not require FM demodulation.

Next, for the second channel at 1599 nm (Fig. 2.22) the reception sensitivities are −14.1 dBm and −9.6 dBm for the positive (□) and the negative (△) weight, respectively. On the other hand, the reception sensitivity for IM input is −15.3 dBm (○). These correspond to reception penalties of 1.2 dB and 5.7 dB for positive (□) and the negative (△) weight, respectively. Nonetheless, the inset in Fig. 2.22 shows that there is no BER floor, and the eye diagram is clearly open. The improved sensitivity for the IM signals compared to the reception of the FM signals is due to the reduced intensity of the ER after FM-to-IM conversion, which results from the moderate contrast in the SOA-REAM comb function. This penalty could be reduced when the functional photodetection is integrated, which will be discussed in the following sub-section.

2.4.4 Synaptic Receptor with Integrated Detector Function

The potential of the SOA-REAM to integrate photodetection functionality is evaluated in the second experimental study, where the device is depicted in Fig. 2.23(a). In this study, its function as a synaptic receptor is compared with FM demodulators assisted by a SoI MRR (Fig. 2.23(b)) and an OTF (Fig. 2.23(c)). For that purpose, the experimental setup is shown in Fig. 2.24.

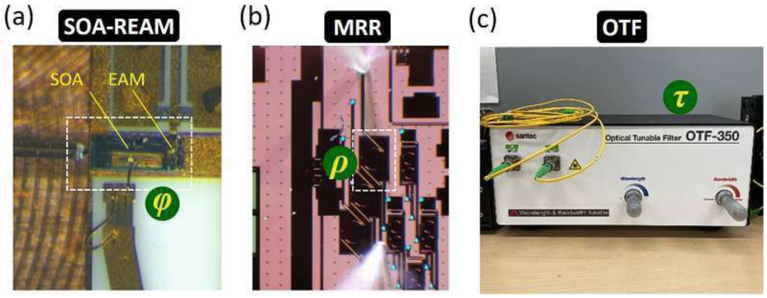

Figure 2.23: FM demodulator devices: (a) SOA-REAM, (b) MRR and (c) OTF [MVS2].

The FM signal is generated at $\lambda = 1585.57$ nm by creating two optical carriers with the first MZM biased for carrier-suppressed double-sideband (CS-DSB) modulation. The two resulting spectral lines at ν_1 and ν_2 are spaced by twice its modulation frequency or $2f_\Delta = 30$ GHz (C). A second 1×2 MZM with complementary output ports modulates both lines with the same pseudo-random bit sequence at a data rate of 2.5 GHz. Both outputs are combined through an optical 25/50 GHz interleaver (IL) to pass only one of the IM lines from each MZM output branch.

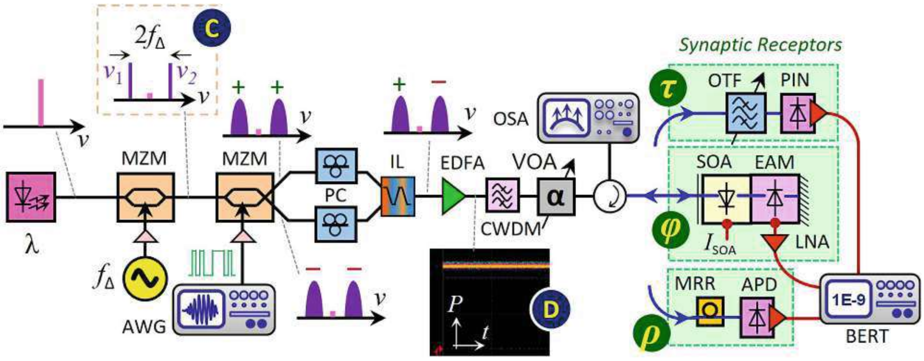

Figure 2.24: Experimental setup for evaluating the SOA-REAM, MRR, and OTF as FM demodulators [MVS2].

Figure 2.25(a) shows a good suppression of 20.4 dB and 19.6 dB between neighbouring odd and even channels of the IL transfer function. This ensures a low crosstalk between the FM tributaries at ν_1 and ν_2. Furthermore, a good optical carrier suppression of 16.8 dB is shown in Fig. 2.25(b)) after CS-DSB modulation. As a result, Fig. 2.25(c) shows a constant-power FM signal, which is obtained with only the tributary at $\nu_2(o)$ and $\nu_1(\varepsilon)$ being found in the lower and upper sideband of the suppressed optical carrier λ, respectively.

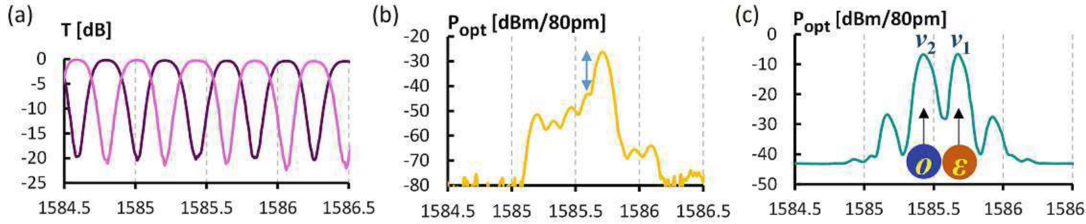

Figure 2.25: (a) Optical interleaver (IL) spectra for odd and even channels, (b) FM signal spectrum, and (c) optical carrier suppression after carrier-suppressed double-sideband (CS-DSB) modulation [MVS2], [MVS3].

It is worth mentioning that the spacing $|\nu_2 - \nu_1| = 2f_\Delta$ is slightly smaller than FSR/2. This is because the spectral configuration of the FM signal, in particular $|\nu_2 - \nu_1|$, is adjusted to the FSR/2 parameter of the demodulating comb function and in order to achieve a good suppression through the even/odd transmission windows of the IL. The particular FM emitter implementation was aimed for a simplified handling of the FM configuration, but it is not a requirement for the proposed synaptic receptor.

The residual IM pattern in the transmitted FM signal has a low ER of 0.26 dB (D in Fig. 2.24). Additionally, the state-of-polarisation at each of the single-mode fiber ports of the IL is aligned with a PC to ensure co-polarized FM tributaries at its output. Subsequently, the signal is amplified with an L-band erbium-doped fiber amplifier (EDFA). Moreover, amplifier spontaneous emission (ASE) peak of the amplifier is filtered with a 1591-nm coarse-wavelength division multiplexing (CWDM) filter.

In the second SOA-REAM based synaptic receptor configuration, the EAM now serves as a photodiode. The detected photocurrent of the EAM is post-amplified with a 50Ω low-noise amplifier (LNA) due to the incompatible EAM interface for low-capacitance co-integration with a transimpedance amplifier (TIA).

For comparison, the SOA-REAM synaptic receptor is replaced with two types of frequency-modulating receivers (Fig. 2.24). The first is a single-polarization MRR, as depicted in Fig. 2.24, followed by an APD/TIA receiver that compensates for the fiber-to-fiber coupling loss of 8.4 dB when passing light through the MRR (ρ). The MRR had been fabricated on a SoI platform with a 220×500 nm^2 waveguide cross section and had a FWHM bandwidth of 14.6 GHz and a FSR of 0.95 nm for its drop response. The second FM receiver implementation is an OTF (Santec OTF-350) with a steep filter edge

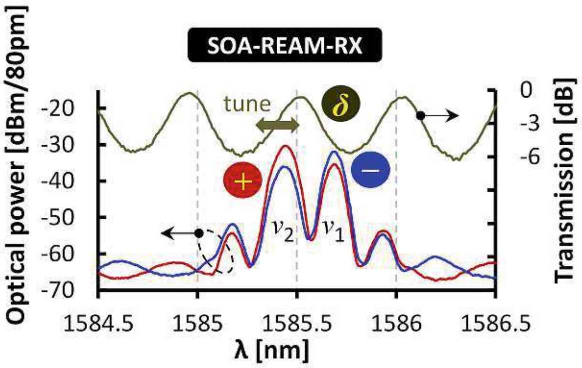

Figure 2.26: Signal spectra for SOA-REAM receptor [MVS2].

of 23.8 dB/0.1 nm and an insertion loss of 2.8 dB, combined with a PIN/TIA receiver (τ in Fig. 2.24).

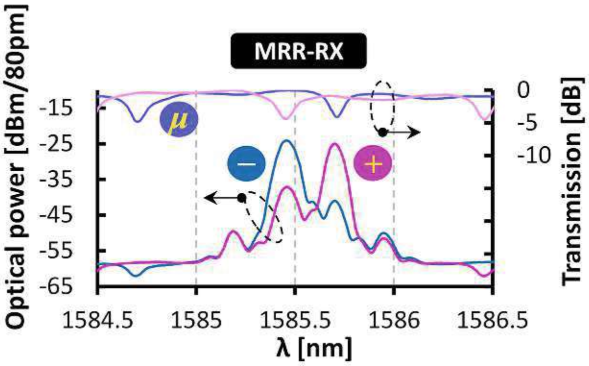

Figure 2.27: Signal spectra for MRR receptor [MVS2].

Figure 2.26 presents the optical signal spectra after FM demodulation through the synaptic receptors for the SOA-REAM, along with the transfer functions of the respective optical elements. The demodulation function of the SOA-REAM (δ) is observed from its reflected (i.e. power-complementary) signal, as shown in Fig. 2.26. It exhibits an extinction of 4.9 dB and can be tuned through the SOA current, the EAM bias or the temperature (Section 2.4.2) to suppress either tributary at ν_1 or ν_2. This results in a positive or negative polarity for the received signal, for which the upper (+) or lower (-) sideband is emphasized, respectively. An extinction of 5.4 dB can be achieved for FM-to-IM demodulation, as it is evident from the difference in peak powers in the upper and lower sidebands of the suppressed optical carrier λ. This extinction has been found to be dependent on the SOA-REAM cavity loss and thus on the EAM bias.

A similar weight setting can be made for the MRR, as depicted in Fig. 2.27. The spectral notch of its pass-through port (μ) is spectrally tuned to reject the corresponding

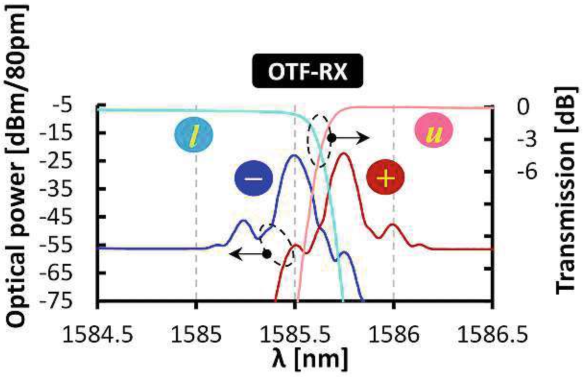

Figure 2.28: Signal spectra for OTF receptor [MVS2].

FM tributary (ν_1, ν_2), leading to an extinction of 15.9 dB for FM demodulation. In the case of the OTF, the filter is tuned to pass either the lower (ι) or upper (u) sideband (Fig. 2.28). The FM demodulation features a high extinction of >32 dB by virtue of the sharp filter roll-off within $2f_\Delta$.

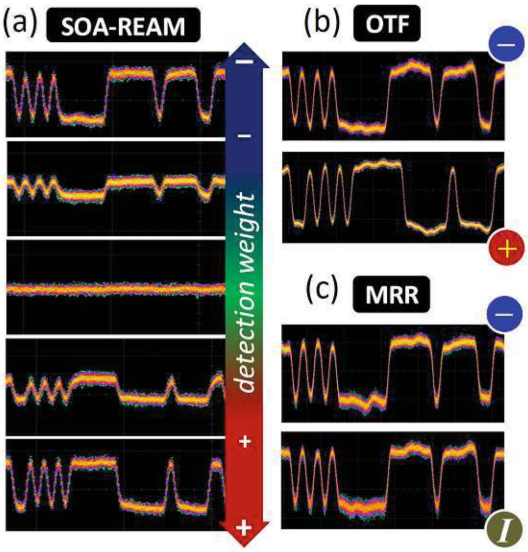

Figure 2.29: Received signal patterns with different synaptic receptors: (a) SOA-REAM, (b) MRR, and (c) OTF [MVS2].

Figure 2.29 reports the detected signals of the frequency-demodulating receivers. A weighting function that is continuous in sign and magnitude is accomplished by shifting the SOA-REAM transmission (a), for which the zero magnitude of weight is shown in the middle. The two insets at the bottom define the positive weight with different magnitudes and the negative signs are depicted at the top. The positive and negative weight can be set in the same manner for the OTF (Fig. 2.29(b)). Similarly, the weight for MRR transmission can be tuned via temperature (Fig. 2.29(c)). Here, the received signal after FM demodulation (–) is compared against the case of having an IM input signal (*I*). The IM signal represents only one FM tributary (ν_2) at the input of the FM demodulator, while the FM demodulation requires to reject the other FM tributary (ν_1).

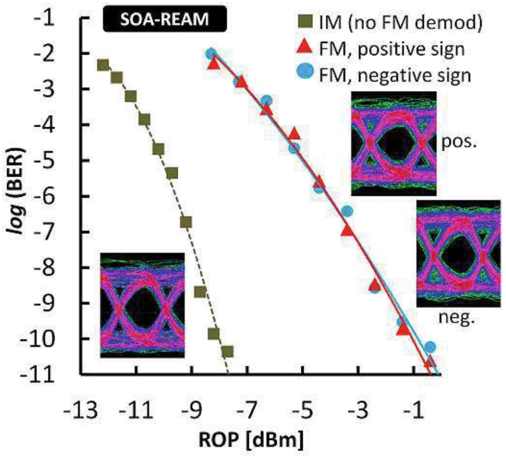

Figure 2.30: BER performances for SOA-REAM receptor [MVS2], [MVS3].

The reception performance of all three synaptic receptors has been evaluated in terms of BER for a pseudo-random data input as function of the ROP. Figure 2.30 shows the BER performance for the SOA-REAM receptor. The reception sensitivity for the FM demodulated signals at a BER of 10^{-10} is $-0.9\,\mathrm{dBm}$ and shows a small difference of 0.2 dB between positive (▲) and negative (●) weights. The penalty with respect to an IM input signal (■) is 7.1 dB, of which 3 dB account for the paired FM spectrum consisting of ν_1 and ν_2 rather than of a single line (ν_2). Another 2.9 dB are attributed to the sensitivity penalty due to a finite demodulation ER (δ) in Fig. 2.26. The rather high ROPs required for the SOA-REAM receptor are attributed to (*i*) the mode-field mismatch between the asymmetric InP waveguide and the tapered single-mode fiber, and (*ii*) the EAM-LNA detector, which performs worse than an EAM-TIA receiver.

For the MRR-assisted receptor with APD/TIA detector (Fig. 2.31), the reception sensitivity is $-14.7\,\mathrm{dBm}$ ($\triangle, \bigcirc$) and the penalty to an IM input ($\square$) is 3.9 dB. Similar as above, the 3 dB is from the suppressed FM sidebands and the 0.9 dB penalty is attributed to the high extinction of the MRR notch used for FM-to-IM demodulation. The high ROP level in combination with the APD receiver are explained by the high fiber-to-fiber coupling loss.

In case of the OTF-based demodulator (Fig. 2.32) with PIN/TIA receiver, a sensitivity of $-10.1\,\mathrm{dBm}$ ($+, \times$) is obtained, whereas the penalty with respect to an IM input (◆)

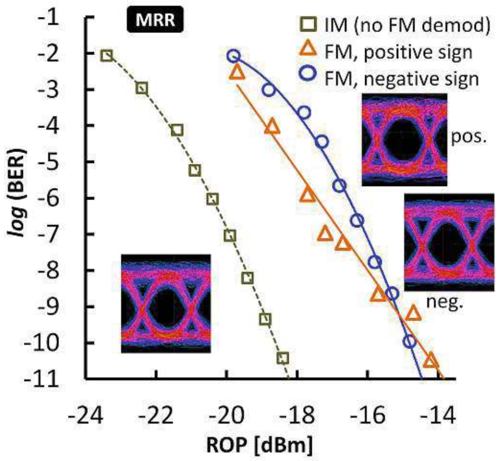

Figure 2.31: BER performances for MRR receptor [MVS2], [MVS3].

is 3.1 dB. This agrees well with the expected 3 dB penalty for an excellent suppression of an FM tributary.

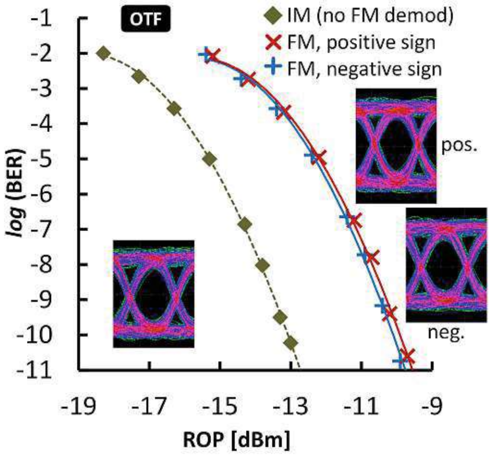

Figure 2.32: BER performances for OTF receptor [MVS2], [MVS3].

2.5 Summary

A new approach of optical linear weighted summation via frequency coding has been presented in this chapter and the results have been published in [MVS1], [MVS2], and [MVS3]. The FM demodulation through a periodic transfer function can facilitate the matrix multiplication between the neural signals and the weights.

The experimental demonstration has confirmed the possibility of frequency coding of neural signals within SOA-REAM synaptic interconnections. Utilizing monolithic integration, the combination of SOA and REAM serves as a colorless synaptic receptor, facilitating the adjustment of both sign and magnitude for linear MAC operations, while simplifying the detector. WDM operation has been validated through the simultaneous demodulation and detection of two spike trains, exhibiting a spectral detuning of 23 nm, consequently enabling a significantly parallel synaptic interconnect.

Additionally, it has been confirmed that a low BER can be observed by integrating the detection functionality with the FM demodulator SOA. A comparison has been done with an MRR- and OTF-assisted FM discriminator, revealing a significant contrast in their transfer functions, suggesting potential improvements in demodulation efficiency for the SOA-REAM receptor. The intensity ER after FM-to-IM conversion was measured at 5.4 dB. Further improvement can be achieved by integrating EAM photoreceiver with an optimized amplifier to TIA to boost the sensitivity of the synaptic receiver.

Integrating the optical elements in ONN can increase the complexity architecture of the NN. The work here confirms that SOA-REAM is practically able to scale up the optical MAC operation. Furthermore, with its colorless feature, the SOA can process multiple synaptic signals with different wavelengths and perform the weighting simultaneously. Therefore, the SOA-REAM offers a breakthrough in the further development of ONN.

Since the linear weighted operation aspect in the optical domain has been demonstrated, the second operation in the ONN requires further investigation. The following chapter will dive into the concept and implementation of an optical nonlinear activation function.

Chapter 3

Nonlinear Activation Function in Photonic Neural Network

The nonlinear activation function is an essential operation after a linear weighted summation in NNs to define a decision boundary in a prediction or a classification problem. Chapter 1 presented some demonstrations of the photonic activation unit with different approaches for various nonlinear functions. In this chapter, another proposal of activation function via frequency coding is introduced. Afterwards, the discussion will continue with the demonstration of photonic ReLU functions by comparing two schemes and their results are reported.

3.1 Concept of Photonic ReLU

Figure 3.1(a) illustrates a photonic neuron within a NN architecture. The first operation is the weight (w_i) multiplication with the neural signal (x_i), resulting in a sum (Σ) in the neuron of the subsequent layer. This linear part has been discussed in Chapter 2, including its demonstration in the optical domain. The second part is the nonlinear operation to perform an activation of the neural summation. Specifically, the study here is devoted to realize a photonic rectified linear unit or ReLU, where the function is illustrated in Fig. 3.1(b). The ReLU function returns to zero if the input has a negative value, but any positive value will pass through, resulting in a linear gradient of 1 (Fig. 3.1(b)). Therefore, the ReLU unit $f(x)$ can be mathematically expressed as

$$f(x) = \begin{cases} x, & x > 0, \\ 0, & x \leq 0. \end{cases} \tag{3.1}$$

where x is the input. This function can be realized in two photonic approaches. Each scheme is discussed in the following.

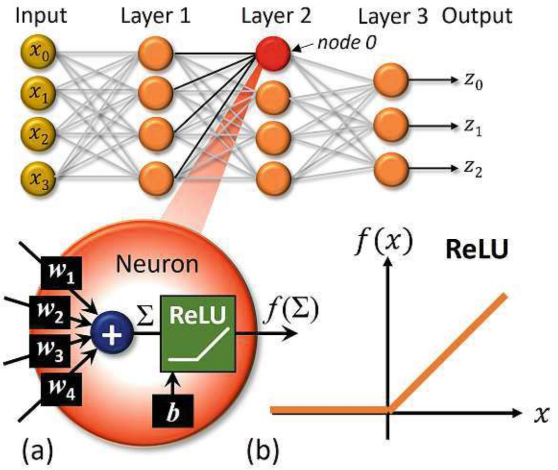

Figure 3.1: Neural network architecture based on multilayer perceptron (MLP) scheme. (a) Functional diagram of a photonic neuron and (b) the schematic of ReLU function [MVS4].

3.1.1 Electro-Optic Slope Filtering

An electro-optic approach refers to a scheme where an electrical-to-optical conversion is required. Here, the DML can be employed to generate the ReLU function, which has been introduced as an FM emitter in Section 2.1.2. Like other semiconductor light sources, stimulated emission can simultaneously occur when a bias current is injected into the DML. However, coherent light is not emitted subsequently, except when the current exceeds a critical value known as a threshold current I_{th}. If the optical input intensity increases and the light excitation is biased above the threshold, the optical output intensity responds linearly to the optical input and is eventually in the lasing state. Therefore, the DML has a light-current ($L - I$) characteristic, as presented in Fig. 3.2(a), which resembles the ReLU function (Fig. 3.1(b)).

To determine how much power (P) is obtained by applying a certain amount of bias current (I), the emission feature of the DML can be characterized via the $L - I$ curve, which is governed by the following equation

$$P = \eta \frac{h\nu}{e}(I - I_{\text{th}}) \tag{3.2}$$

where η is the quantum efficiency of the laser, $h = 6.626 \times 10^{-34}$ J·s is a Plank constant, ν is the laser operating frequency, and $e = 1.6 \times 10^{-19}$ C is the electron charge.
The direct electro-optic modulation property of DML is utilized to perform the photonic ReLU activation function. The neural weighted sum signal is sent via a RF signal generator to the DML (Σ_{el}), as depicted in Fig. 3.2(b). Afterwards, a bias current (I) is added

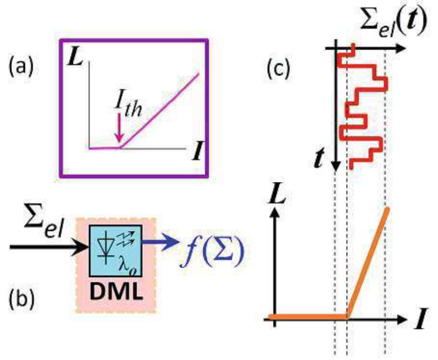

Figure 3.2: (a) The light-current ($L - I$) characteristic of DML, (b) the electro-optic ReLU scheme using DML, and (c) the alignment of the RF weighted sum data (Σ_{el}) to the light-current curve of DML to perform ReLU operation can be controlled via the bias current [MVS4], [MVS5], [MVS6].

to Σ_{el} to align the neural signal to the $L - I$ curve of DML, as illustrated in Fig. 3.2(c). The sub-threshold bias current ($I < I_{\text{th}}$) is required, as the Σ_{el} below the threshold will be rectified. Therefore, by tuning the bias current of the RF signal, the photonic ReLU function $f(\Sigma)$ can be configured.

3.1.2 Photonic Slope Filtering of Frequency-Coded Synapses

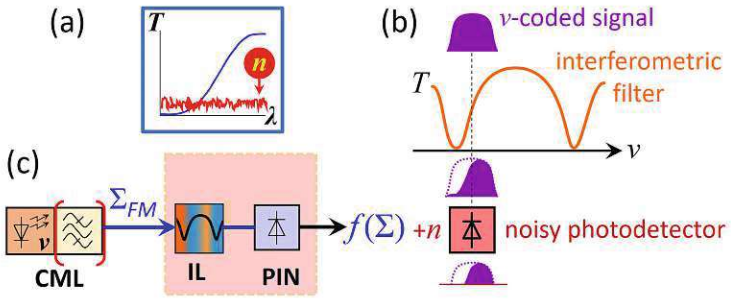

Figure 3.3: (a) Illustration of a linear transfer function of an interferometric filter with the noise floor n of a photodetector, (b) a photonic ReLU scheme, and (c) the concept of photonic ReLU by using a combination of CML+IL and the subsequent PIN photodiode [MVS5], [MVS6].

The second approach passes a frequency-domain representation of the optical signal to the quadrature point of an interferometric filter, where the linear transmission function is illustrated in Fig. 3.3(a). It can be noticed that the transmission is similar to the

ReLU function. However, it has a curvilinear feature, which can be inappropriate for performing rectification because the cut-off threshold for negative input value is not well defined. Nonetheless, the output signal will eventually be received by a photodetector and the proper rectification can be done through its inherent clipping feature deriving from its finite sensitivity.

This comes from the fact that a photodetector is subject to noise that limits the sensitivity of the receivers due to several effects [111]. Through its load resistance, the photodetector contributes with so-called thermal noise. This thermal noise results from the temperature fluctuation of the charged carriers. Additionally, current fluctuation is present in course of the so-called shot noise, which results from the discrete nature of photons that follow a Poissonian statistic. The received photocurrent on the detector is likely prone to typical thermal and shot noise [112], thus giving rise to a noise floor (red; n), as depicted in Fig. 3.3(a). The n provides a cut-off point to realize the rectification operation. Thus, the combination of the interferometric filter and the photodetector provides the transfer function analogous to the ReLU function (Fig. 3.1(b)).

To quantify the sensitivity of a photodetector due to the noise, a noise equivalent power (NEP) can be measured, which defines the incident power on the detector that results in a signal-to-noise ratio (SNR) of one at the outputs at a given data signal rate and noise bandwidth [111]. The NEP is expressed in the $W/\sqrt{Hz}$ unit. Since NEP defines the lowest optical signal that can be detected, it is beneficial to have a detector with a small NEP value, which directly corresponds to a lower noise floor, meaning a higher sensitivity.

In order to perform the activation function via the interferometric filter, the neural information is coded to the frequency domain. The neural data $s(t)$ is translated to the optical frequency ν, according to the following mathematical expression

$$s_\nu(\nu) = \nu_\lambda + \Delta\nu \cdot s(t) \tag{3.3}$$

where ν_λ is the optical carrier frequency of the frequency-coded signal and $\Delta\nu$ the peak-to-peak frequency deviation, which is applied during information coding. The signal s_ν is delivered through the quadrature point of the interferometric filter, which is illustrated in Fig. 3.3(b).

An optical interleaver (IL) has a similar transfer function as the interferometric filter. It has been used widely in optical communication applications to combine two sets of WDM signals from odd and even channels and interleave them into a composite signal. The transmission of IL from the even channel can be defined through a periodic function as

$$T(\nu) = \frac{1}{2}\left[1 - \cos\left(2\pi\delta T(\nu - \nu_o)\right)\right] \tag{3.4}$$

where δT is the delay parameter that defines the FSR and ν_o is the offset of the transfer function to the FSR.

Therefore, photonic ReLU via IL and photodetector can be performed by feeding in an FM signal, meaning that a FM transmitter is required precedingly. The superiority of a CML as a FM emitter has been discussed in Section 2.1.2.2 and can be used as a good companion with the IL and the noisy photodetector. Therefore, the FM signal of the weighted sum Σ_{FM} can be generated via the CML and is further sent to the transfer function of the IL, in combination with the subsequent photodetector, to process ReLU $f(\Sigma)$ (Fig. 3.3(c)). The rectification can be achieved by tuning the wavelength of the CML relative to the IL slope function. In the next section, the analytical model of these approaches are reported.

It is important to mention that even though a photonic approach involves optoelectronics, the next chapter will show that a translation between the optical and electrical domains between the layers of the ONN does not require complex interfaces.

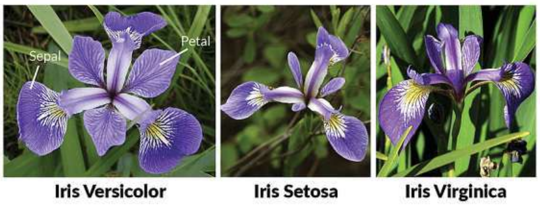

Figure 3.4: Iris flower as classification problem example [113].

One of the well-known classification problems is an Iris flower data set, introduced by a statistician and simultaneously biologist, Ronald Fisher, in 1936 to analyze a linear discriminant function [114]. The Iris dataset includes data from 150 flowers. It is distributed from three classes of the flower: Iris setosa, Iris versicolor, and Iris virginica, as shown in Fig. 3.4. Each flower sample has four features, consisting of sepal (length and width) and petal (length and width), where all are defined in centimeter (cm) and collected by Edgar Anderson [115].

For that purpose, Fig. 3.1 introduces a multilayer perceptron (MLP) architecture with three hidden layers besides the input and output layers. In each hidden layer, there are four nodes, except layer 3, which has only three nodes, same amount as the classes of the Iris flower. On the other hand, the four nodes in the input layer are as many as the features of the flower, where each node (x_i) in the input layer represents one flower feature.

As a benchmark, the Iris flower prediction problem has been trained offline with 100 epochs using the Adam optimization [116] by Pham [117]. There are 55 trainable parameters comprising 20, 20, and 15 parameters for hidden layers 1, 2, and 3, respectively. It should be pointed out that the activation function does not occupy the parameters. Additionally, a standard partition of 80% for training and 20% for testing has been implemented in the training. In addition, a learning rate of 10^{-3} and a batch size of 20 have been used in the training optimization. The results of the training becomes the digital NN (DNN) and is used as a benchmark for the photonic neuron demonstration.

Therefore, each neuron in NN is weighted with an integer (w_i) as the result of the training. Afterwards, MAC operation is performed, and bias (b) is applied to their sum (Σ). Then, the nonlinear activation function (f) is utilized, and the output is sent to the subsequent layer. A Softmax function (z_i) is applied at the output layer to determine the prediction probability. For performance evaluation, the outcome of photonic ReLU demonstrations will be compared to the ideal DNN.

3.2 Analytical Model

Figure 3.5(a) shows the transmission model T (dashed-line) based on Eq. (3.4) and the 25/50 GHz IL transmission (line). As can be seen, the model fits to the measured transmission. In order to perform the rectification operation, the linear function can be estimated at its quadrature point, at $T = -3\,\text{dB}$. Therefore, the model in Eq. (3.4) can be approximated to the following function

$$L(\nu) = \alpha + \beta[\nu - \nu(T = 0)], \tag{3.5}$$

where

$$\alpha = T(\nu(T = 0)) = \frac{1}{2}, \tag{3.6}$$

and

$$\beta = \frac{d}{d\nu}T(\nu(T = 0)) = \pi\delta T \tag{3.7}$$

Figure 3.5(b) presents the approximation based on Eq. (3.5)-(3.7), including the transmission model T (dashed-line). This IL function selectively passes the spectral components of the neural signal s_ν. In the scenario illustrated in Fig. 3.3(b), the filter suppresses lower frequencies without rectification. To achieve rectification, the subsequent photodetector after IL introduces a noise background n, which is effectively responsible as the drop window (D in Fig. 3.5(b)) for the linear function L. This clips the neural signal, as defined by the bias b_ν for the spectral alignment of s_ν. Therefore, the neural signal in the higher frequencies and above D will pass through. Eventually, this scheme provides the ReLU function for photonic neuron implementation.

As mentioned earlier, the ROP to the photodetector will determine the threshold point n. In order to explain the influence, two cases are shown in Fig. 3.5(c), where two schemes of ROP are defined. If the ROP is too large, meaning that the sensitivity of the photodetector is high, then the threshold n' will be low. As the consequence, a nonlinear characteristic is shown below −3 dB transmission of the IL, as pointed out by e_{IL}. It will be eventually lead to an incorrect ReLU operation due to the nonlinearity above n'.

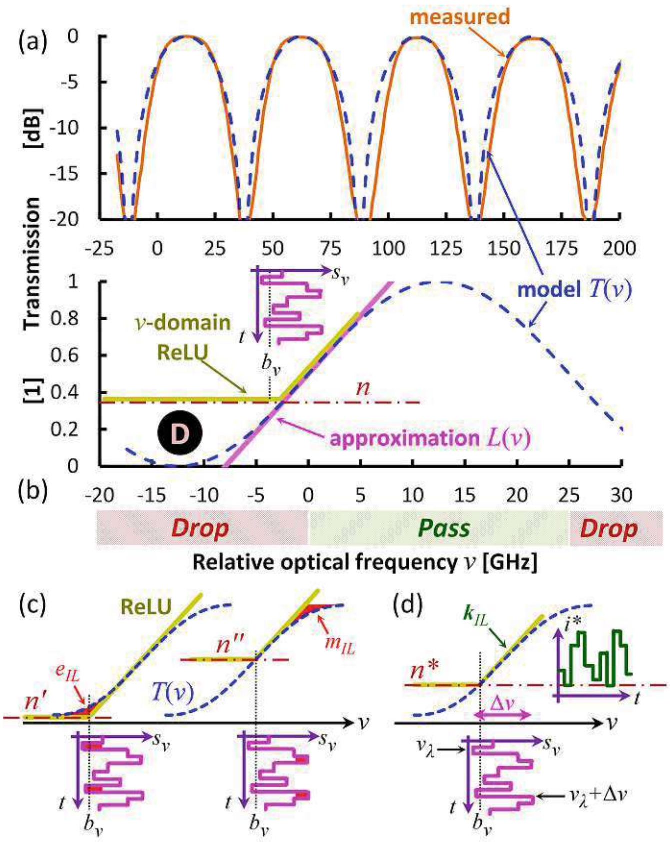

Figure 3.5: (a) The model T (dashed-line) and measured (line) transmission of an optical 25/50 GHz interleaver as a function of a relative optical frequency. (b) A linear approximation L_ν (line) is around the quadrature point of the model (dashed-line). The noise background n from a photodetector is clipping the signal, which separates the pass and drop regions of the IL. (c) Sub-optimal noise thresholds where nonlinear distortion for the neural signal appears at the extinction e_{IL} and maximum transmission m_{IL} of the IL. (d) Optimal threshold n^* yielding a rectified photocurrent i^* defined by the frequency deviation $\Delta\nu$, the ReLU bias b_ν, and the frequency-to-intensity conversion k_{IL} of the IL [MVS6].

On the contrary, if the ROP is too low, then the sensitivity of the photodetector is moderate. As a result, the threshold n'' will be too high and the neural data signal will end up above the -3 dB value of the IL slope. At the maximum of the IL transmission, a nonlinear response as indicated by m_{IL} will affect the performance of the ReLU function as well.

Therefore, there is a trade-off that requires an optimization to avoid these nonlinear responses. In order to minimize the error ξ in the ReLU function due to the two sub-optimal cases above, the ROP has to be optimized, which is given by the deviation from its linear slope according to the following equation

$$\xi = \int_{\nu_\lambda}^{\nu_\lambda+\Delta\nu} |T(\nu) - L(\nu)|\, d\nu \tag{3.8}$$

and as indicated by the red areas at e_{IL} and m_{IL} in Fig. 3.5(c).

The ideal threshold condition n^* is sketched in Fig. 3.5(d), which is neither not too low nor too high as in Fig. 3.5(c). The photocurrent of the processed neural data i^* (after frequency-to-intensity conversion k_{IL} at its linear slope), is correctly rectified without additional error. In addition, it requires a careful match of the frequency deviation $\Delta\nu$ to the FSR of the IL to preserve the neural signal s_ν within the linear region $L(\nu)$. It is also important to appropriately scale $\Delta\nu$ as it cannot be made arbitrarily small to ensure linear operation because this would result in a degraded signal-to-noise ratio. Therefore, a trade-off between non-linear distortion and excess noise must be made.

The experimental demonstration of photonic ReLU will be presented in the next section and their results are discussed afterwards.

3.3 Experimental Evaluation

3.3.1 Response Characterisation of the ReLU Function

Two characterisations of photonic ReLU are presented, which are based on the electro-optic of DML and the frequency coding using IL and photodetector. Figure 3.6 shows the laser characteristic of a DML (Yuopto YDBK) exhibiting the output power against the applied bias current. The model (dashed line) is illustrated based on Eq. (3.2) and fits to the measured DML laser (solid line) at 1554 nm with $\eta = 0.218$ and $I_{\mathrm{th}} = 6.5$ mA. For the demonstration of the ReLU function, the rectification can be applied at the sub-threshold current regime by tuning the bias current accordingly.

Based on the analytical model in Section 3.2, characterization is performed to experimentally identify the ReLU function response in the case of the frequency coding approach. Figure 3.7 shows the experimental setup to perform this characterization. A vector network analyzer (VNA) drives an optical inphase/quadrature modulator (IQM) through a wideband RF 90° hybrid to generate an optical single-sideband frequency sweep. A 25/50 GHz IL (Optoplex IL-C0SBFAS004) with a slope of $k_{\mathrm{IL}} = 5.43$ dB/GHz is then inserted between IQM and PIN photoreceiver (HP 11982A) with a NEP of 30 pW/ $\sqrt{Hz}$.

Figure 3.7 presents the corresponding transfer function when spectrally allocating the optical source at the beginning of IL transmission window. At the same time, the launch power is adjusted with an attenuator (VOA) to accomplish clipping of the filtered response. For an average ROP below −10.5 dB, a suppression towards lower frequencies

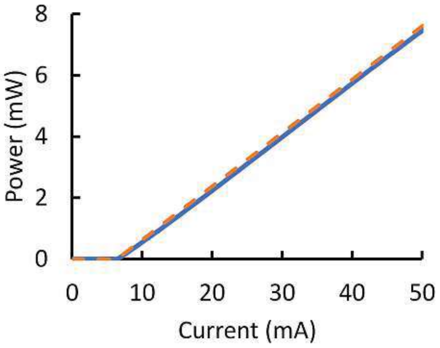

Figure 3.6: The model (dashed line) and measured (solid line) $L-I$ curve of a DML (Yuopto YDBK).

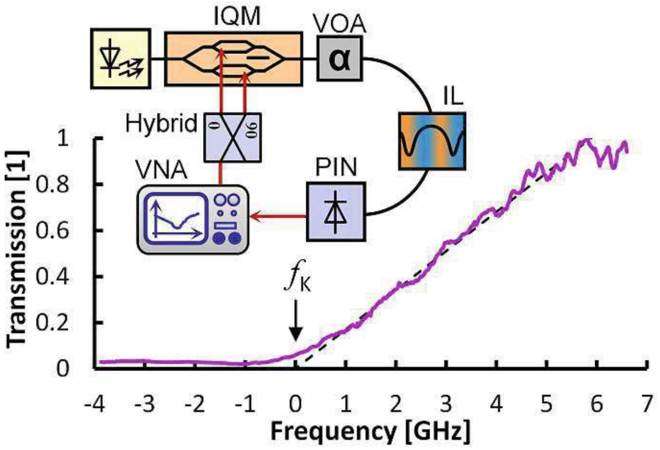

Figure 3.7: Characterisation demonstration of ReLU response [MVS6].

$f < f_K$ can be seen, with $f_K = 0$ GHz. A linear slope is obtained above f_K, which together with the suppressed region resembles a ReLU function. Therefore, the IL and photodetector will be used in the demonstration below to perform the photonic ReLU function.

3.3.2 Demonstration of ReLU Function for Photonic Neural Network

Figure 3.8 shows the experimental setup to evaluate the performance of the photonic ReLU function. The evaluation is conducted based on a neuron in layer 2, node 0, as denoted in the MLP architecture of Fig. 3.1. The weighted sum data $\Sigma w_i s_i(t)$ is sent through an AWG at 1 Gb/s data rate, which drives a CML operating at 1600.6 nm. It is important to mention that the sum $\Sigma w_i s_i(t)$ data has been selected without being

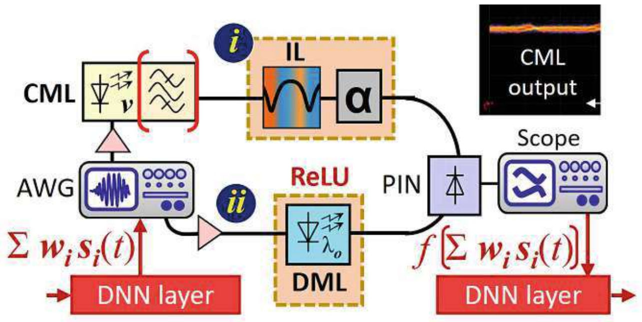

Figure 3.8: Experimental setup to evaluate the performance of ReLU unit based on (i) an optical interleaver (IL) and a PIN photodetector, in combination with a CML, and (ii) an electro-optic approach by employing a DML [MVS4], [MVS5], [MVS6].

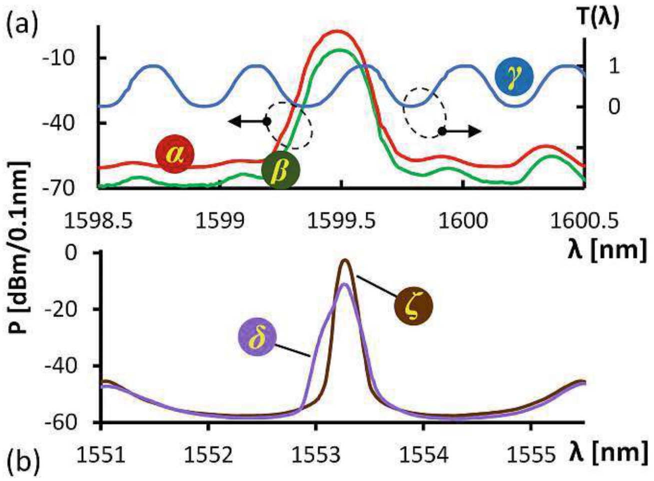

Figure 3.9: (a) Optical spectra of CML+IL based on ReLU implementation before (α) and after (β) the IL, including the IL transmission (γ). (b) DML output spectrum for bias points above (ζ) and below (δ) the threshold current I_{th} [MVS4], [MVS5], [MVS6].

processed by the activation function in DNN. The CML is used as a synaptic emitter to generate the FM signal of weighted sum data Σ_{FM}, where its OSR filter has been adjusted to a neutral position to prevent FM demodulation (Section 2.1.2.2). The inset in Fig. 3.8 depicts the output of the CML with an intensity ER of 0.3 dB. The FM signal is processed by the 25/50 GHz optical IL (Fig. 3.8(i)) and the subsequent photodetector to perform the ReLU activation function, which inherently converts the frequency-to-intensity modulated signal.

Figure 3.9(a) displays the optical spectra of CML before (α) and after (β) the IL, when the wavelength λ of CML is aligned to the IL linear slope, along with the transfer function of the even IL channel (γ). Afterwards, the optical signal output after the

ReLU is received by a PIN photodetector and digitized through a real-time oscilloscope (Fig. 3.8). As performance comparison, the combination of CML+IL is replaced by a DML operating at 1554 nm, performing the electro-optic ReLU with a threshold current I_{th} of 6.5 mA (Fig. 3.8(ii)). In this investigation, two distinct bias currents are used, where Fig. 3.9(b) shows the optical spectra of DML with the sub-threshold current of 5.5 mA (δ) and above I_{th} of 11.2 mA (ζ), which is in the linear regime of the $L-I$ curve (Fig. 3.6).

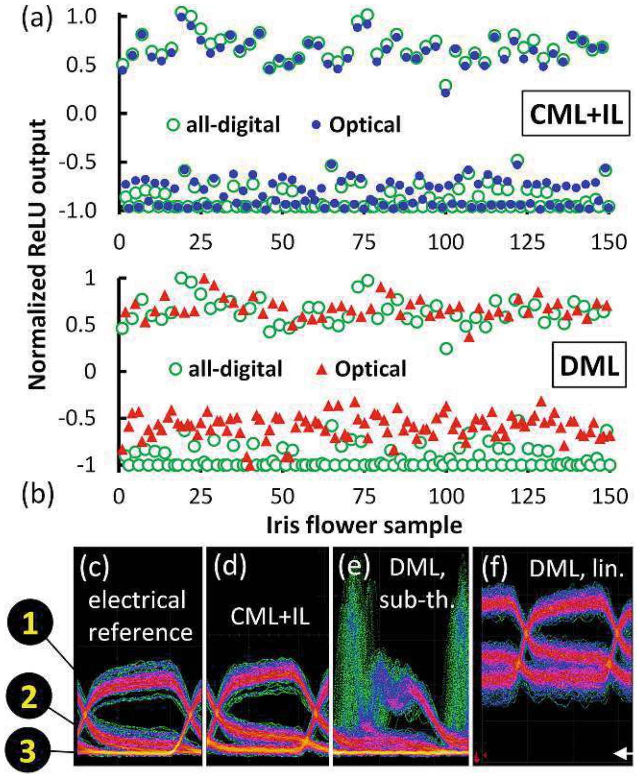

Figure 3.10: Normalized optical signal at the output of ReLU for digitally computed NN (○) and photonic ReLU implementations based on (a) CML+IL (●) and (b) a DML (▲). The eye diagrams at the bottom are presented as the output ReLU for the case of (c) all-DNN, and photonic ReLU using (d) CML+IL and (e) DML with bias current below and (f) above the threshold current [MVS4], [MVS5], [MVS6].

Figure 3.10(a) describes the normalized optical signal with CML+IL (●) at the output of the ReLU for 150 Iris flower data samples, being compared to all-digital NN (○). It is visible that the photonic ReLU process with the CML+IL is in a good agreement with the digitally-computed NN. This can also be observed from the equivalent eye diagrams (Fig. 3.10) at the output ReLU between the reference DNN (c) and the CML+IL (d).

On the other hand, the ReLU operation with the DML (▲) reveals a more significant error compared to the DNN-based ReLU (○), as reported in Fig. 3.10(b), by applying a sub-threshold current to the DML to rectify the neural signal. The error arises from a gain switching effect, as the eye diagram (e) shows that the DML emission is pulsed. This also corresponds to the broadened optical spectrum δ with the biased current below the I_{th} in Fig. 3.9(b). To mitigate this effect, the bias is increased above I_{th}, where the ReLU occurs in the linear region of $L - I$. However, the subsequent eye diagram (f) exposes an inaccurate rectification, as the eye is shifted to the top. This indicates the unsuitability of the DML to perform a ReLU operation.

The error when performing the photonic ReLU by an implementation based on CML+IL and DML are summarized in the histogram of Fig. 3.11. Here, the error ε is introduced as an absolute difference between the normalized optical output signal and the all-digital NN. The histogram shows that the error by implementing a CML+IL based ReLU (κ) is small with an average error $\bar{\varepsilon} = 0.14$ ($3\sigma = 0.36$). In contrast, the error is larger by using the DML, which is distributed in a range of $0 < \varepsilon < 0.7$. The results assure that the CML+IL combination can be used to perform photonic ReLU with a low error.

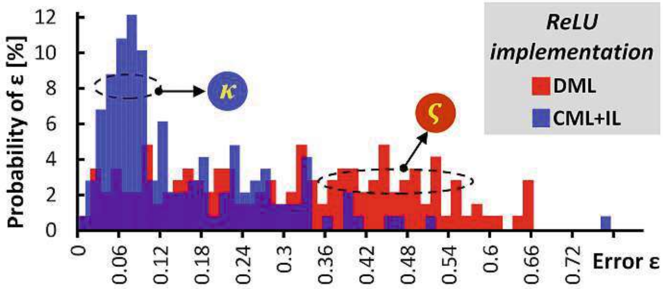

Figure 3.11: Error histogram of photonic ReLU operation based on CML+IL (κ), which is lower than using DML implementations (ς) [MVS4], [MVS5], [MVS6].

In addition, the output of the photonic ReLU is fed back to the DNN. This corresponds to propagating the measured optical neural signal to the output layer of the MLP architecture (Fig. 3.1), aiming to calculate the accuracy of Iris flower prediction. The accuracy of 90% and 62% is achieved for the CML+IL and DML implementations, respectively. It is noteworthy that the accuracy is achieved by sending one flower every nanosecond. In the case of offline training, the DNN reveals 93% accuracy for the Iris flower prediction. This highlights the potential of CML+IL and the subsequent photodetector scheme as the favourable photonic ReLU, leaving just a 3% penalty of accuracy compared to the digitally-computed NN.

It is worth noting that the DNN did not consider any hardware or optical component imperfections when making its training considerations. The DNN training is directly transferred and applied to the optical layer without taking into account the actual subcomponent implementations in the demonstrated photonic neuron.

3.4 Summary

It has been demonstrated that photonic ReLU can be achieved based on frequency-domain processing involving combination of CML+IL with a subsequent photodetector. The results have been published in [MVS4], [MVS5], and [MVS6]. Furthermore, the results show that a lower error is achievable with CML+IL rather than a DML implementation, which is prone to undersired lasing artifacts in the operation regime required to perform non-linear electro-optic signal processing. Using the CML+IL tandem with noise background of a photodiode shows that a non-linear electro-optic conversion can be avoided when performing the ReLU function. In addition, a small penalty of 3% can be accomplished by using this scheme, compared to the offline training of the DNN.

The full implementation of a photonic ReLU including an analog linear weighted summation is the next step. This frequency coding based on CML+IL and the noise-clipped photodetector will be employed for further investigation in developing a functionally complete photonic neuron, which will be discussed in the next chapter.

It is important to mention that even though a photonic approach involves optoelectronics, the next chapter will show as well that a translation between the optical and electrical domains between the layers of the ONN does not necessarily require complex interfaces. The next chapter will therefore further introduce a translucent concatenation between the linear and nonlinear operations in photonic domain, where no electrical amplification is required.

Chapter 4

Photonic Neural Network

Two essential ingredients in ONNs, the analog linear weighted summation and the photonic activation function based on frequency coding, have been introduced and demonstrated in Chapters 2 and 3, respectively. Therefore, the next step is to combine all the functions to realize a functional photonic neuron. This chapter will show that multiple neural sub-circuits can be collapsed on the same photonic ReLU hardware. This signifies the prospect of integrating the weighting functionality within the photonic ReLU for a simplified ONN scheme. Afterwards, the possibility of simultaneously performing the weighting function and ReLU unit will be investigated. This work reveals that combining the linear and nonlinear functions is challenging because it requires an intermediate optical-to-electrical-to-optical (o/e/o) conversion. Therefore, a direct-drive approach will be presented, which eliminates an electrical amplification in this o/e/o process. This will open a new path towards an efficient interface in ONNs. At the end of this chapter, the possibility of integrating the proposed photonic neuron scheme on a chip-scale implementation will be discussed.

4.1 Methodology: Heterodyne Optical Detection

In order to analyze the possibility of weight polarity assignment via the photonic ReLU and to characterize the FM signal at the DBR laser output in the following experimental demonstration, a coherent optical detection will be employed to restore the full information on the optical complex amplitude. Optical coherent detection is a technique to measure the phase and the amplitude of a transmitted signal at the receiver [118]. Coherent detection retrieves the phase information of the signal which cannot be recovered by a direct detection technique. Figure 4.1(a) illustrates the schematic of heterodyne coherent detection where the incoming signal at frequency ω_S is combined with a locally generated signal from a laser oscillator (LO) laser at frequency ω_L. Balanced photodetectors is used to suppress the direct current (DC) component and maximize the beat between the two signals [119]. The total current at the photodetectors can be defined as

$$I(t) = R|A_{\mathrm{S}}\mathrm{e}^{-i\omega_{\mathrm{S}}t} + A_{\mathrm{L}}\mathrm{e}^{-i\omega_{\mathrm{L}}t}|^2 \tag{4.1}$$

where R is the photodetector responsivity, $A_{\mathrm{S}} = a_{\mathrm{S}}\mathrm{e}^{i\phi_S}$ and $A_{\mathrm{L}} = a_{\mathrm{L}}\mathrm{e}^{i\phi_L}$ are the complex amplitude of the transmitted signal and the LO signal laser, respectively. By defining $P_{\mathrm{S}} = |A_{\mathrm{S}}|^2$ and $P_{\mathrm{L}} = |A_{\mathrm{L}}|^2$ as the signal power and LO power, respectively, the total current can be written as

$$I(t) = R\left[P_{\mathrm{S}} + P_{\mathrm{L}} + 2\sqrt{P_{\mathrm{S}}P_{\mathrm{L}}}\cos\left(\omega_{\mathrm{IF}}t + \phi_{\mathrm{S}} - \phi_{\mathrm{L}}\right)\right] \tag{4.2}$$

where $\omega_{\mathrm{IF}} = |\omega_{\mathrm{S}} - \omega_{\mathrm{L}}|$ is an intermediate frequency (IF), ϕ_{S} and ϕ_{L} are the phase of the signal and LO laser, respectively. Equation 4.2 describes that the signal phase ϕ_{S} can be retrieved as long as ϕ_{L} is constant. Additionally, Fig. 4.1(b) shows the initial position of the LO frequency ω_{LO} and the signal frequency ω_{S}. By setting the LO frequency close to the signal frequency, the IF spectrum can be achieved, as illustrated in Fig. 4.1(c). This term is known as down-conversion, where the IF signal accurately represents the original signal, meaning that the amplitude and phase of the signal is preserved. The phase and the amplitude of the signal are recovered electronically at an IF ~1 GHz [118]. When $\omega_{\mathrm{IF}} = 0$, then it corresponds to a so-called homodyne detection [119].

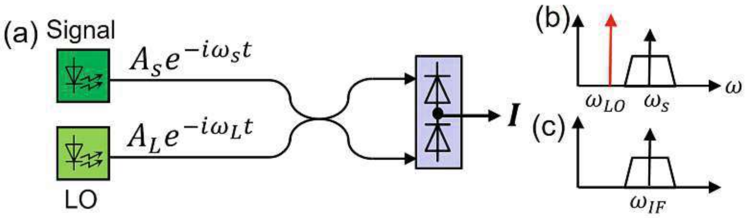

Figure 4.1: (a) Illustration of a coherent detection and spectra of (b) the optical signal and (b) the down-converted signal.

In the following, the experimental demonstration of the photonic neuron will be reported and the results will be discussed.

4.2 Photonic Neuron with Weighting and Nonlinear Activation Functions

This section will demonstrate a photonic neuron consisted of the analog linear weighted summation and the photonic ReLU for two approaches. The first involves collapsing multiple neural sub-circuits on the proposed ReLU to set the polarity of the synaptic weights. This scheme explores the potential for utilizing a shared photonic hardware across multiple neuromorphic sub-circuits. The second approach is a simultaneous operation of weighting summation and the ReLU unit. This task will investigate whether the two operations can be performed in parallel.

4.2.1 Collapsing Multiple Neural Sub-Circuits onto a Shared Photonic ReLU

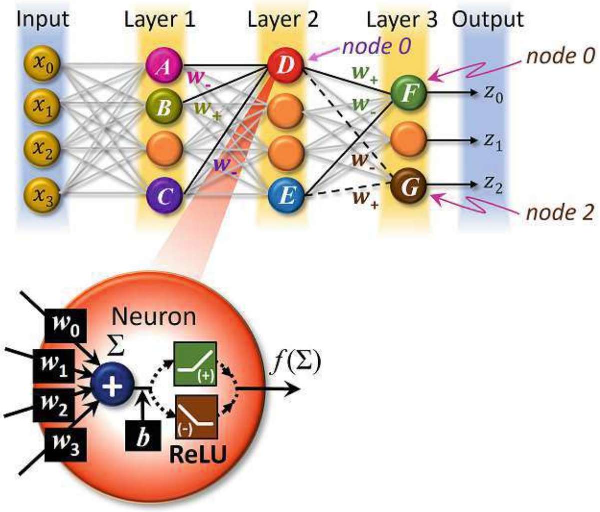

Figure 4.2: Neural network architecture based on the multilayer perceptron scheme of a photonic neuron, comprising the linear weighted summation (Σ) and the photonic ReLU [MVS6].

Figure 4.2 shows the NN architecture to evaluate a photonic neuron with a shared IL. Two neurons under consideration are denoted in layer 2, nodes D and E. The weights applied to the neurons are based on the offline training or the digital NN (DNN) and can have different polarities and magnitudes, depending on the neuron under investigation in the subsequent layer. Based on the DNN, the weighted sum of nodes D and E, including the activation function afterwards, will result in layer 3, nodes F and G, as illustrated in Fig. 4.2, with alternating signs (w_+, w_-) for the weight.

Figure 4.3 reports the experimental setup, where the IL is the shared ReLU for two neural sub-circuits. All synaptic emitters utilize the same IL. In this evaluation, four CMLs are used as FM synaptic emitters; two of them are allocated in the C-band (1535.2 and 1535.55 nm), and the other two are in the L-Band (1599.65 and 1600.5 nm). Both C-band CMLs define the synaptic emission of two nodes in layer 2 (D and E in Fig. 3.1). The corresponding wideband PIN photodetector (sub-circuit 1; Fig. 4.3) receives the output representing the neuron in layer 3, node 2 (G) after the ReLU activation. Likewise, both L-band CMLs define the same emitters (D and E), and the output

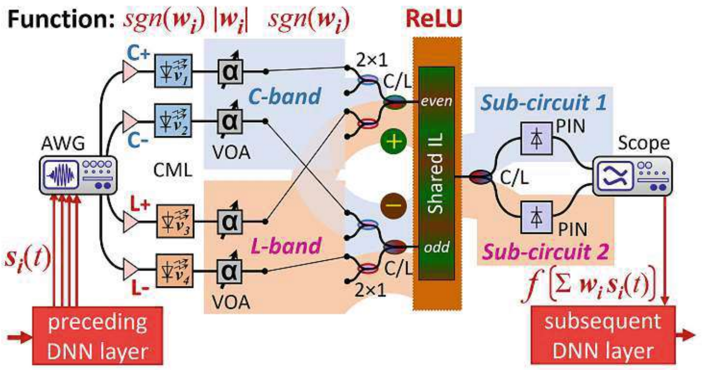

Figure 4.3: Experimental setup to evaluate an ONN based on frequency-domain ReLU with multiple neural sub-circuits using a common ReLU [MVS5], [MVS6].

result at the L-band PIN receiver (sub-circuit 2) is devoted to the neuron at layer 3, node 0 (F).

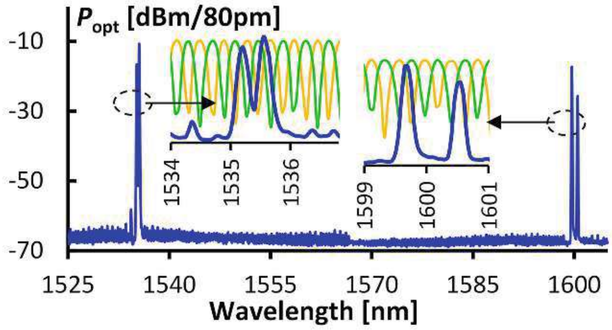

Figure 4.4: Optical spectra of four CMLs in the C- and L-band after shared IL. The insets highlight the spectral allocation of the CMLs to the even and odd channels of the IL for assigning the weight polarity [MVS5], [MVS6].

The data $s_i(t)$ of the corresponding neurons in layer 2 from DNN (without being weighted digitally) are sent to the AWG to drive the CMLs at a signal rate of 1 Gb/s (Fig. 4.3). The weight assignment, $sgn(w_i)$, is accomplished through a spectral alignment of the frequency-coded signals of CMLs output to the periodic transfer function of the IL with respect to its positive or negative slopes (Section 2.1.1). Figure 4.4 describes the optical spectra of all CMLs after the shared IL, including the insets showing the transmission function of IL from the even and odd channels. The insets indicate the CMLs spectral allocation to the interferometric IL slopes (rising or falling) following the corresponding synaptic weight assignment (positive or negative).

Moreover, Fig. 4.2 shows that a VOA is positioned in each output of the CML for setting the weight magnitude $|w_i|$, taking advantage of its frequency-agnostic feature.

The $|w_i|$ is configured based on the following formula

$$|w_i| = 10\log_{10}|w_{\mathrm{DNN}}|^2 = 20\log_{10}|w_{\mathrm{DNN}}| \tag{4.3}$$

where w_{DNN} is the weight value based on the DNN training, a rational number in the range between [-1,1]. Equation 4.3 comes from the fact that the w_{DNN} is comparable to power in the electrical domain ($w_{\mathrm{DNN}} \sim P_{\mathrm{electrical}}$). On the other hand, the power in optical domain is proportional to the square of the $P_{\mathrm{electrical}}$ ($P_{\mathrm{optical}} \sim P^2_{\mathrm{electrical}}$) and $|w_i| = 10\log_{10} P_{\mathrm{optical}}$. Therefore, the $|w_i|$ yields to the Eq. 4.3 in the optical space.

It is worth mentioning that the $|w_i|$ does not necessarily need to be modified through the VOA. Section 5.4 will explain that $|w_i|$ can be adjusted through the drive of the FM emitter, which is more practical.

Next, 1×2 splitters are utilized to facilitate the necessary weight combinations and connections toward the even (+) and odd (-) inputs of the shared IL (Fig. 4.3). C/L waveband splitters are used afterwards to separate the two ONN sub-circuits. The signal is then received by the PIN photodetector in either waveband, and the output is digitized through a real-time oscilloscope to calculate the accuracy of the Iris flower classification case.

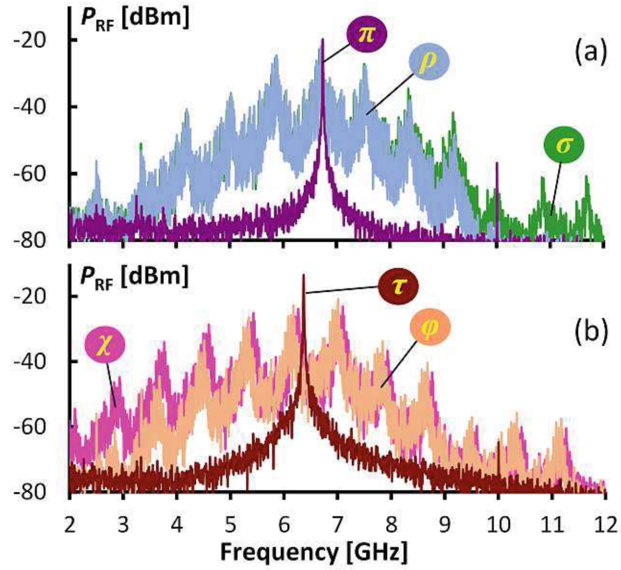

Figure 4.5: Heterodyned spectra for (a) negatively and (b) positively weighted neural signal [MVS5], [MVS6].

Figure 4.5 shows the synaptic weight setting using the photonic ReLU, and the frequency-coded ReLU operation, employing a CML with a chirp parameter of 2.9. The spectra present the continuous-wave beat notes (π, τ) for the unmodulated C-band CML and a reference lasers, resulting in an IF of ~6.6 GHz. In addition, the RF spectra depict the heterodyned neural signals before (σ, χ) and after (ρ, ϕ) passing the signal to the shared IL for a negative (Fig. 4.5(a)) and a positive (Fig. 4.5(b)) polarity assignment. In the case of a negative sign, proper rectification occurs as the falling slope of the IL (ρ) suppresses the higher frequencies of the optical frequency-coded signal, as shown in Fig. 4.5(a). In contrast, for the positive sign setting, the rising slope of the IL suppresses the lower frequencies of the neural signal (ϕ in Fig. 4.5(b)).

Figure 4.6 presents the error results for the two ReLU sub-circuits with the shared IL in a histogram, visualized similarly to Fig. 3.11. The average error $\overline{\varepsilon}$ remains consistently low at 0.1 for both neurons in layer 3, node 0 (μ, L-band) and node 2 (ν, C-band). The accuracy prediction is performed by feeding back the optical signal output data to be processed by the remaining layers of the DNN.

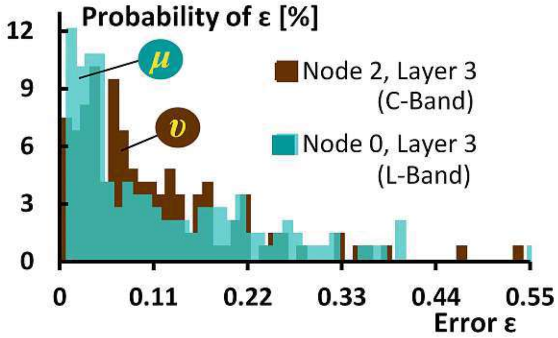

Figure 4.6: Error histogram for the photonic NN output of node 0 (μ) and node 2 (ν) in layer 3, after performing optically weighted sum and ReLU function [MVS5], [MVS6].

Figure 4.7 summarizes the accuracy of both neurons as a function of $\overline{\varepsilon}$. For both neurons at layer 3, an accuracy of 92% is achieved at the 1 Gb/s data rate of the experiment. This result closely matches the accuracy of 93% achieved by the DNN. As mentioned in Chapter 3, the result is obtained without adjusting the NN training procedure to the involved optical elements under investigation.

Afterwards, the influence of the error on the accuracy is investigated. For that purpose, noise is added to the experimental data to simulate the degradation of the neural signal. With an increase in $\overline{\varepsilon}$ due to an artificial noise degradation, a significant drop in accuracy for node 2 (▲) is observed. In contradiction, the accuracy for node 0 (●) decreases only slightly. This discrepancy can be attributed to the varying noise tolerance levels of the involved neurons, suggesting the need for precise optical implementations for particularly noise-sensitive neurons.

Furthermore, Fig. 4.8 illustrates the relationship between $\overline{\varepsilon}$ and the neural data rates for the noise-sensitive node 2 at layer 3. The error is evaluated for experiments at different data rates (1, 2.5, and 5 Gb/s). It can be seen that between the data rates of 1 and 5 Gb/s, the $\overline{\varepsilon}$ increases slightly from 0.10 to 0.12. The increased error leads to a drop in accuracy to 90% at 5 Gb/s.

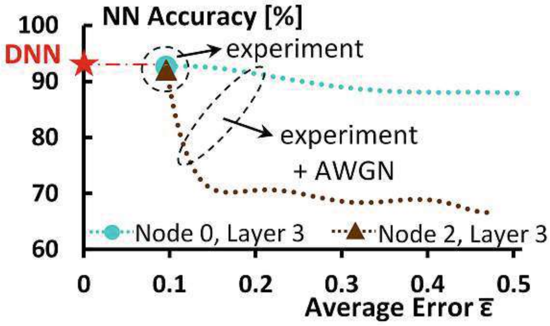

Figure 4.7: The NN accuracy for Iris flower classification at 1 Gb/s of node 0 and node 2 in layer 3, after performing optically weighted sum and ReLU function.

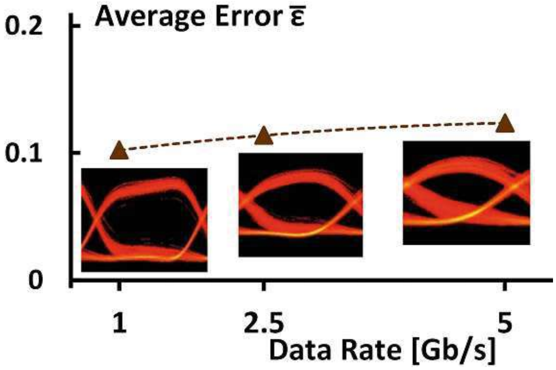

Figure 4.8: Average error $\bar{\epsilon}$ as a function of the data rate for node 2 in layer 3 [MVS5], [MVS6].

The following sub-section will discuss a different approach of a simultaneous scaling of the linear weighted summation and the photonic ReLU function.

4.2.2 Concurrent Linear Weighted Summation and ReLU

The second implementation scenario focuses on evaluating the possibility of jointly performing the ReLU function and the weight setting. The experimental setup is presented in Fig. 4.9 and involves a single neural circuit with three synaptic emitters. The weights are assigned to the outputs of the neurons in layer 1 (*A*, *B*, *C* in Fig. 4.2) and then summed up. After the ReLU activation operation, the output results in the neuron of node 0, layer 2 (*D*). Two CMLs in the L-band represent the emitters of the two nodes in layer 1 (*A* and *C*), while one CML in the C-band is used for the node *B*.

Similar to the first scenario, the $s_i(t)$ data is sent to the CMLs of the corresponding neurons in layer 1 at 1 Gb/s through the AWG. The signal in the C-band (node *B*) has a positive polarity (w_+), with the weight w_{DNN} is 0.484. In contrast, the other two

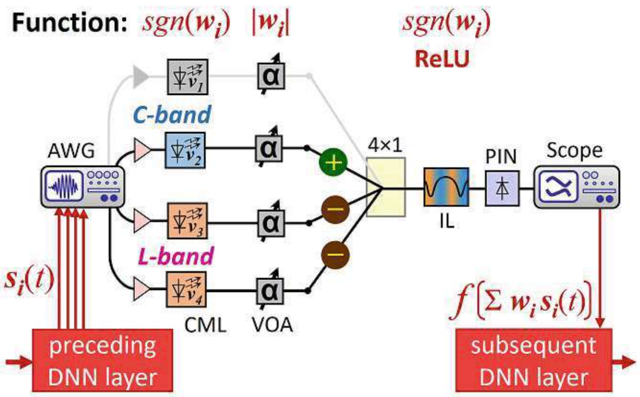

Figure 4.9: Experimental setup to evaluate an ONN based on frequency-domain ReLU with simultaneous weighting and ReLU activation function when scaling up the number of neural inputs [MVS6].

neural data streams in the L-band (A and C) have a negative sign (w_{-}), as illustrated in Fig. 4.9, with the w_{DNN} of -0.574 and -0.145. The three synaptic signals are combined with a colorless 4×1 splitter and then sent to the IL to execute the ReLU activation function. Finally, the signal is detected by a PIN receiver, and the accuracy of the ONN is determined again after digitizing the neural signal output.

An average error $\overline{\varepsilon}$ of 0.2 is obtained in this demonstration, which is higher than the result in the first scenario (Fig. 4.6). Despite the moderate error, the optical output signal is fed back to the DNN architecture. An accuracy of 91% is still achievable for the Iris flower classification case. However, this could be attributed to a noise-insensitive neuron configuration, given the relatively high $\overline{\varepsilon}$ value.

The discrepancy can be explained by the fact that integrating the weight polarity setting with the ReLU function does not accurately mimic the original function of the neuron, as the ReLU is not applied to the weighted sum. Therefore, an accurate neural signal processing in a scaled-up ONN would necessitate a clear separation between the linear and the nonlinear sub-circuits within the photonic neuron, which is evident from the demonstration of the first scenario in Section 4.2.1.

Nevertheless, the other reason of the high error can come from the fact that one of the w_{DNN} is smaller compared to the other two. The sum is then more or less similar to the large weight that dominates the linear weighted summation. Therefore, if there is a dominating weight, the entire linear and nonlinear circuits can be simplified at the cost of a small additional error.

In the next section, it will be demonstrated that a complex separation between the linear and nonlinear operations is not necessary.

4.3 Synthesizing Neural Network Layers with Linear and Nonlinear Functions

As discussed above, the linear and nonlinear operations can not be integrated easily in the photonic neuron because it is typically involving o/e/o conversion at the demarcation point. This section will discuss the idea towards all-optical end-to-end processing in the ONN.

4.3.1 Concatenating Linear and Nonlinear Functions

The frequency coding of neural signals is the main concept proposed in this thesis for the analog weighted addition and the photonic activation ReLU. In the linear operation, the summation takes place in a wideband photodetector, taking advantage of its wavelength-agnostic characteristic (Chapter 2), which converts the received optical weighted sum signal into a photocurrent representation. On the other hand, an optical FM transmitter is required for the ReLU activation circuit. The transmitter sends the frequency-coded weighted sum signal, which is subsequently processed by the photonic ReLU unit. This indicates that the o/e/o process is necessary at the demarcation point between $\sum$ and ReLU in the photonic neuron (Fig. 4.2), in order to translate the photocurrent of the weighted sum into a frequency domain representation for the subsequent activation operation. On the other hand, the o/e/o conversion can become a challenge when scaling up the ONN. Especially, it can lead to a power consumption overhead when an electrical amplifier is involved for this purpose.

In certain scenarios, the laser such a DBR laser can be directly driven by the photocurrent from the detector, eliminating the need for an additional RF amplifier. This comes from the fact that the DBR laser can provide a large conversion efficiency for translating the signal from the amplitude to the frequency space. It has been demonstrated that a DBR laser can have a broad tuning range of 40 nm [120]. Furthermore, it has been reported recently that a DBR laser can have a translation efficiency up to 47.6 GHz/mA [MVS7]. As a result, the output of the photodetector is sufficient to drive the DBR laser. The transmitter optical sub-assemblies (TOSAs)-packaged DBR lasers facilitate the translation of electrical signals $s(t)$ into optical frequency-domain representations $s_\nu(t)$, as defined in Eq. 3.3. Therefore, the combination of a PIN photodiode and a DBR laser forms a transparent intensity-to-frequency converter, where the detected photocurrent is translated into an optical signal based on the frequency coding.

Figure 4.10 shows the tuning characteristics of the DBR lasers. When the Bragg section of the DBR laser is injected by current, the effective refractive index n_{eff} is getting smaller which subsequently decreases the Bragg wavelength (Eq. 2.8), shifted by $\Delta\lambda_{\text{B}}$. Therefore, the reflectivity peak will shift to a lower wavelength. On the other hand, the wavelengths that meet the requirement of the phase condition shifts to lower wavelength with the displacement of $\Delta\lambda$. If the $\Delta\lambda_B > \Delta\lambda$, which is usually the case, the mode jumps to the adjacent lower wavelength that satisfies the phase condition, resulting to mode hopping [121], as presented in solid line of Fig. 4.10.

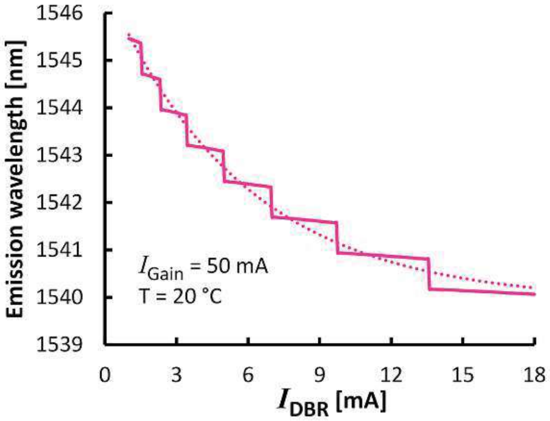

Figure 4.10: Characterisation of a DBR laser [MVS8].

Apart from the discrete jumps due to the DBR modes, the tuning slope is governed as in the following

$$\lambda = \lambda_0 + \lambda_S \cdot \exp(-I_{DBR}/I_\tau) \tag{4.4}$$

where λ_0 is the baseline wavelength, λ_S is the peak-to-peak swing for tuning, λ_{DBR} is the bias current, and I_τ is the FM efficiency current for tuning. The dotted line in Fig. 4.10 is the tuning slope of the DBR laser used in the experiment, fitted according to Eq. 4.4 by choosing $\lambda_0 = 1539.9$ nm, $\lambda_S = 6.7$ nm, $\lambda_{DBR} = 1.5$ mA, and $I_\tau = 5.8$ mA.

It shall be noted that a low signal rate of 100 Mbaud will be chosen in this demonstration because the electro-optic modulation response of the DBR laser in this work (3-dB bandwidth) rolls off at 270 MHz [MVS7]. The following sub-section will demonstrate the feasibility of the direct-drive scheme in the photonic neuron.

4.3.2 Translucent Concatenation of Weighted Sum and Nonlinear Activation

Figure 4.11 shows the experimental setup to demonstrate the possibility of omitting the electrical amplifier in the o/e/o conversion between the linear and nonlinear operations in the photonic neuron. Here, a direct-drive scheme is introduced where a DBR laser is directly driven by the photocurrent, which is generated by a preceding photodiode.

In this demonstration, the neuron in node 0, layer 2 (D) in the DNN is considered again, based on the MLP architecture in Fig. 4.2. The weighted sum signal $\Sigma w_i s_i(t)$ of the node (without being processed by the activation function digitally) is sent with

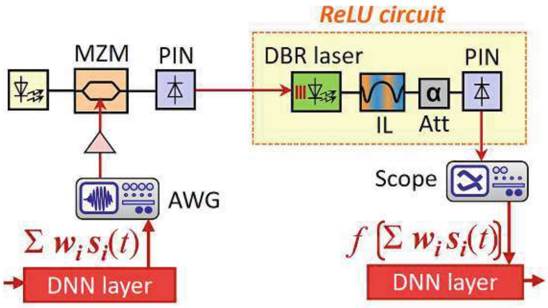

Figure 4.11: Concept and experimental setup for an o/e/o conversion between the linear and the nonlinear part of the photonic neuron, employing a DBR laser directly driven by a PIN photodiode current. The same photonic ReLU scheme is used to evaluate the translucent concatenation scheme [MVS6].

100 Mb/s information rate to a MZM modulator, which modulates an optical signal at 1550 nm, as depicted in Fig. 4.11. This optical weighted sum Σ_{FM} is received by a PIN photodetector, which subsequently drives the DBR laser in the ReLU circuit to perform the activation operation.

The DBR laser in the ReLU circuit translates the photocurrent to the optical frequency coded weighted sum Σ_{FM} as mathematically described in Eq. 3.3. Next, similar to the experiment in Chapter 3, the wavelength of the DBR laser is aligned to the IL transmission by tuning the bias current of the gain section of the DBR laser. Then, the ReLU function is performed by the IL and the subsequent PIN photodetector. The output signal is eventually captured by the scope for further offline evaluation.

The efficiency of the DBR laser for the frequency translation process of a neural input is characterized and reported in Fig. 4.12 with a 100 Mbaud signal rate. For this purpose, the output of the DBR laser was directly captured using a high-resolution optical spectrum analyzer. Additionally, it was heterodyned with a reference laser to acquire its FM spectrum at a down-conversion frequency of ~5.5 GHz. Figure 4.12 shows a ROP of −7 dBm to the PIN photodetector, which subsequently drives the DBR section, results in an optical frequency-translated neural signal with a deviation at the DBR laser output of $\Delta\nu = 4$ GHz. This provides large spectral expansion by translating from the amplitude to the frequency space without requiring an electronic amplifier. The outstanding spectral tuning efficiency of the DBR laser drives this prominent ratio between the frequency deviation and the data symbol rate.

Additionally, Fig. 4.12 demonstrates the spectral alignment of the DBR laser with respect to the IL transmission function. The spectrogram shows that the FM signal at the DBR laser output exhibits a deviation sufficiently wide to perform the rectification. This effect corresponds to the (partial) suppression of levels 2 and 3 of the neural data (as a reference, see Fig. 3.10(c)) after passing the ReLU optical function.

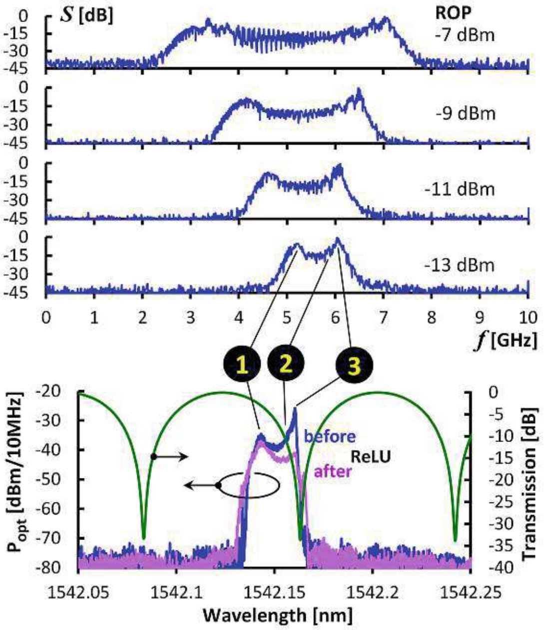

Figure 4.12: Spectrogram of the DBR laser output and spectral alignment of the DBR laser to the IL transfer function to realize a photonic ReLU function [MVS6].

Figure 4.13 compares an optical output signal performed by ReLU (red trace) and a digitally performed ReLU (yellow). The optical signal and the digital computation reference are in a good agreement, even though a bandwidth limitation in the optical signal is apparent. As mentioned previously, the restricted electro-optic 3 dB bandwidth of the DBR section is responsible for this bandwidth limitation in the optical signal. Nonetheless, the levels of neural signal align well with the digital levels at the optimal decision sampling point. This is evident from the error histogram in Fig. 4.13, which reveals an average error $\overline{\varepsilon}$ of 0.1 using the proposed direct-drive DBR scheme as the FM transmitter. The $\overline{\varepsilon}$ is marginally lower than the one obtained with CML+IL based scheme, which was evaluated in Section 3.3.

Moreover, the optical ReLU output signal of node 0, layer 2 is fed back to the MLP architecture in the digital domain and propagated forwardly to the output layer, aiming to determine the accuracy of the Iris flower problem. The accuracy of 92% is obtained, thus resulting in a 1% penalty in comparison to the DNN. This result demonstrates that an efficient o/e/o conversion between two photonic signal processing circuits can be realized, as the electrical signal conditioning can be avoided.

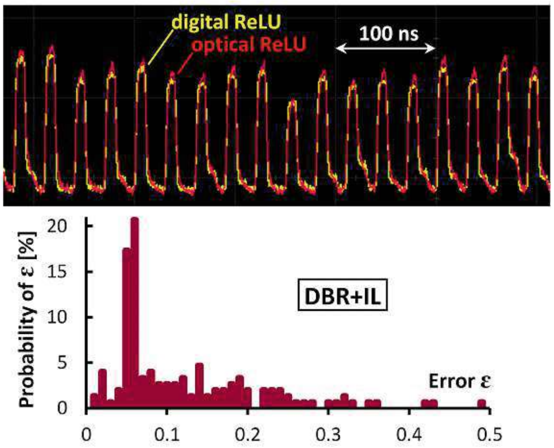

Figure 4.13: ReLU output from the optical and all-digital NN implementation, and corresponding error histogram for the optically implemented ReLU function [MVS6].

4.4 Chip-Scale Integration

Integrated photonics technology has developed over a decade since 2008 and matured rapidly [122]. The opportunity is now broadly open to integrate the optical elements for neuromorphic photonic purposes on-chip because it can significantly scale down the space occupied by the neurons. This will drive a fast photonic neural network field development for wide-range applications.

Silicon photonics is the forefront technology and has become the foundation for neuromorphic computing research. Some efforts have been made to realize silicon neuromorphic photonic circuits [36], [50], and [123]. It has been demonstrated that an SoI-based ONN can reach information rates of tens of GHz [124].

However, it has been known that silicon cannot facilitate light emission purely due to its indirect semiconductor bandgap. Over the years, this has become the bottleneck in combining active and passive components on a single silicon-based PIC. Therefore, most applications have been performed on a hybrid integration approach with external optical power supply. Many have worked on external lasers based on III-V semiconductors with high-quality performances [125]. However, the coupling loss can be a factor that degrades the performance of the integrated ONN since it requires a precise alignment between the SoI waveguide PIC and the laser, which further incurs cost due to complex chip-to-chip or chip-to-die assembly. On the other hand, it is possible to facilitate hybrid integration on-chip, employing a hetero-integration approach [126], [127]. However, a unique fabrication process is necessary, while massive batch production is difficult and expensive.

Besides silicon, Indium Phosphite (InP) is a promising PIC platform since it features a direct bandgap that permits a monolithic integration of optical source, signal manipulation and detection. It has been demonstrated that InP-based emitters have outstanding features [79], [128]. Moreover, it is possible to fabricate other main optical building blocks with InP [129], such as amplifier [130], modulator [131], and detector [132], [133]. Therefore, it is feasible to integrate all the active and passive optical components on a single InP-based circuit. Several demonstrations of InP-based PIC have been evaluated for different applications, such as quantum key distribution [134], light detection and ranging (LIDAR) [135], [136], and photonic neural network [56]. Therefore, an InP-based integrated ONN is a promising solution and likely to yield favourable results.

In order to implement the presented frequency-domain based photonic neural network in this thesis, it is necessary to find the equivalent building blocks in the PIC. From the light emitter side, directly modulated lasers such as CML and DBR lasers have been employed here as FM transmitters. On the other hand, integrated InP DBR laser has been theoretically studied and experimentally demonstrated [79], [137]. Furthermore, the InP DBR laser has been available as a design file so-called a process design kit (PDK) for fabrication in some foundries [138], including also the PIN photodiode. The PDK includes foundry-specific data files for component and chip design, which is owned by foundries [139], and it is usually used as a basic design in a mass production, known as a multi project wafer (MPW) run [140].

Even though the aformentioned InP DBR laser requires an additional modulator to generate FM signal, a recent study on directly modulated laser (DML) on InP membrane platform has been proposed, allowing a possibility to be fabricated in the future [141]. The FM-to-IM conversion for the weight assignment and activation function, the equivalent interferometric filters, such as delay interferometer or optical interleaver, are not easily fabricated. The most common fabricated filter is a DBR grating and now is available on MPW run too [142].

Nonetheless, some efforts are being made to fabricate an integrated InP delay interferometer [84] and an AWG-based integrated optical interleaver [143]. Even though the layouts are complex, monolithic integrated ONN architectures have been developed and are continually progressing to address the challenges. Eventually, the photonic neuron proposed in this study will be feasible to be implemented on-chip to support further scaling of the number of neurons that will functionally power the ONN.

4.5 Summary

It has been demonstrated that the function of a photonic neuron can be performed with a low error by combining linear weighted summation and the ReLU in the optical domain. These demonstrations have been reported in [MVS5] and [MVS6]. The simultaneous weight assignment can be done using the interferometric filter IL, which additionally features a photonic ReLU unit. Sharing the photonic ReLU hardware facilitates a simplified and cost-optimized layout for the ONN. A low penalty of 1% in accuracy compared to the DNN for Iris flower classification supports the practicability of implementing the scheme in a larger neural network computation.

In addition, it has been evaluated that the scalable solution of the simultaneous weighting and ReLU operation shows a moderate error. These results show the necessity of a clear separation between the linear and nonlinear activation functions. Therefore, the translucent concatenation concept has been introduced to solve the typically inefficient o/e/o process between optical circuits. This translucent scheme is accomplished through a tandem of PIN photoreceiver and DBR laser, where the detected photocurrent drives the DBR laser directly, without necessitating RF circuits for signal conditioning. The low penalty of 1% compared to the DNN convinces that the scheme can enable an energy-efficient interface in ONN implementations where multiple neural sub-circuits are to be concatenated.

A future refinement can be made by replacing the DBR laser under investigation with one featuring a larger electro-optic modulation bandwidth. Another research has proven that DBR lasers can be compatible with modulation up to 10 GHz [144]. Furthermore, the on-chip scaling of the proposed ONN architecture is feasible when considering compact integrated optical elements, as they are available through state-of-play photonic integration platforms.

In the next chapter, a photonic processing unit will be presented, which performs neuromorphic processing and other signal processing applications. Furthermore, the photonic neuron based on switch-spaced architecture be compared with the one using wavelength-routed architecture, which is potential for the scaling of ONN in the future advancement.

Chapter 5

Photonic Processing Unit for Photonic Signal Processing

Digital signal processing (DSP) plays an important role in optical communication networks for a wide range of applications, such as dispersion compensation, signal modulation, fiber transmission, and signal detection. However, most of them are performed in the digital domain to carry out signal modulation, signal shaping, and signal recovery. On the other hand, recent advancements in components and integrated circuits in the field of photonics allow signal processing to be performed directly at the optical layer. This chapter presents a multi-purpose photonic processing unit, the so-called ΦPU, which performs functions in optical communication networks, such as modulation format decoding and neuromorphic processing. The motivation for an ΦPU will be floored at the beginning of this chapter. Afterwards, the concept of each of those functions will be introduced. Next, the experimental demonstration of ΦPU based on space-switched architecture will be presented and the results will be reported. Additionally, a photonic neuron is demonstrated and discussed in a wavelength-routed neural interconnect that adopts the direct-drive scheme as a comparison to the space-switched approach.

5.1 Motivation

Chapter 1 has described the explosion of the internet and its effect on global energy consumption. On the other hand, the current core network capacity is reaching a bottleneck since it cannot sustain the demand anymore in the future, resulting in the so-called capacity crunch [145]. This pushes some improvements in the systems, including the use of DSP in the optical communication network to expand the transmission capacity [146]. The role of the DSP is especially essential in employing higher-order modulations using four degrees of freedom in the optical domain [phase, amplitude, optical frequency, and polarization] to encode and access information [147]. This paves the way for a high-capacity transmission and enhanced spectral efficiency. However, since the DSP has been implemented in an optical network, it involves optical-to-electrical (o/e) and electrical-to-optical (e/o) translations between the photonic and digital domains [148]. As a result, this adds to inefficient energy consumption, which can be avoided if all processes are executed in one domain only.

On the contrary, advancements in photonic technology have allowed optical signal processing to be performed efficiently without the need for electronic sampling and opto-electronic conversions, owing to the vast photonic bandwidth and low energy consumption [149]. Some demonstrations have shown signal processing directly performed at the optical layer, such as optical beamforming networks for remote radio heads [150], a reconfigurable photonic signal processor [151], or neuromorphic circuits in receivers serving fiber non-linearity compensation [152], and radio source separation [34]. Nevertheless, each modulation or processing technique is conventionally underpinned by a particular system dedicated to a specific objective. On the other hand, if all the signal processing functions are collapsed in a single system, the optical elements can be re-used or shared for a simplified infrastructure.

Towards the purpose of the general optical processing node, this chapter presents a photonic processing unit (ΦPU) to solve optical communication challenges directly at the optical domain without o/e or e/o conversions. The typical processing challenges covered in this thesis involve signal decoding (duobinary and PAM4) and neuromorphic processing. The following section will discuss the concept of each signal processing task in the optical domain.

5.2 Concept of Reconfigurable Photonic Architecture

The proposed ΦPU is shown in Fig. 5.1, which is set up from an 8 × 8 space-switch. It includes transmitters (Tx) and receivers (Rx) on the left- and right-hand side, respectively. The main components of the ΦPU are built on a set of opto-electronics converters and passive optical elements. This includes a set of DBR lasers as an efficient FM translator of the input signals on the transmitter side. The receiver side consists of optical photodetectors that convert the signals back to the electrical domain. The switch is additionally supported by the colorless splitter for signal distribution or combination within the ΦPU and a set of periodic filters to process the signal, such as decoding or rectification operation. The input signal is supplied by a waveform generator, and the output signal is digitized for performance evaluation.

The ΦPU configuration is flexible, allowing multiple processing circuits to be collapsed over the same optical switch matrix. Furthermore, it is reconfigurable because the components within the ΦPU core can be re-used for different signal processing chal-

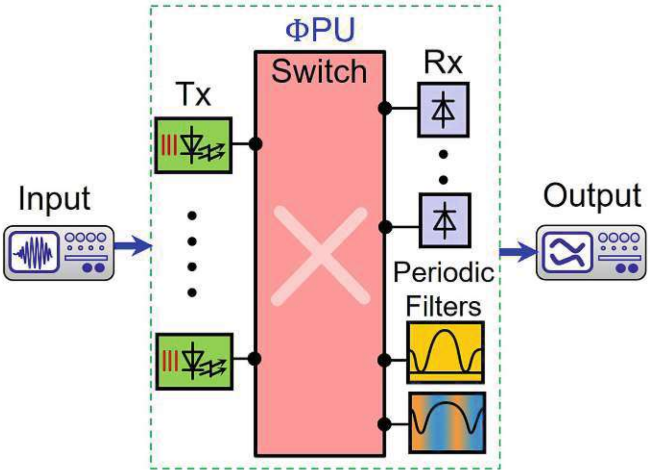

Figure 5.1: The concept of a photonic processing unit (ΦPU) which includes transmitters, receivers, and periodic filters for signal processing.

lenges. The following section will explain the concept of each optical signal processing task under consideration.

5.3 Concept of Photonic Signal Processing

In this section, the concepts of several signal processings are described and their implementations in the optical domain are explained.

5.3.1 Duobinary Decoding

The duobinary format has been widely used in optical communication networks to increase channel capacity over a narrow-bandwidth channel [153]. Besides high spectral efficiency, the duobinary signal is more resilient to chromatic dispersion at higher bit rates than non-return-to-zero (NRZ) signals [154]. Duobinary code represents symbols –1, 0, and +1, as depicted in Fig. 5.2. In terms of intensity, the symbols –1 and +1 correspond to an “on” level, while the symbol 0 correlates to an “off” stage.

In the conventional analog duobinary system, the transceiver corresponds to a duobinary signal generator, while the receiver acts as a full-wave rectifier to convert the duobinary to binary signal, as introduced in [155]. However, the recent applications with a high Gbaud data rate signal processing demand a broadband full-wave rectifier with a flat response over the entire data bandwidth, which is challenging to practically realize [156]. Some digital duobinary decoding alternatives have been proposed [157], [158]. However, e/o and o/e converters are required in those schemes, together with costly and energy-hungry analog-to-digital converters.

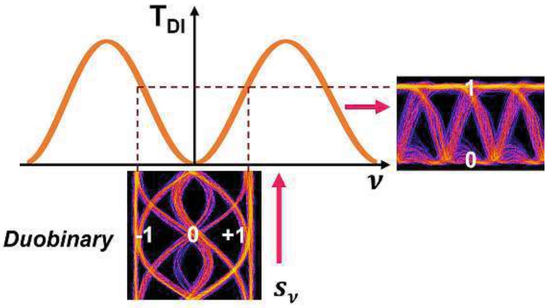

Figure 5.2: Concept of photonic duobinary decoding in the ΦPU using a delay interferometer (DI).

To perform full-wave rectification in the optical domain, the duobinary decoding can be performed in the frequency domain by utilizing the cosine-shaped transfer function of an interferometric filter, as shown in Fig. 5.2. Given its wide FSR, full-wave rectification can occur around the quadrature points of the filter.

The frequency-coded representation of the duobinary signal is essential to employ this optical scheme. Therefore, the electrical duobinary data signal is precedingly converted to its optical frequency domain representation via an efficient DBR laser, which has been discussed in Section 4.3.1. One of the interferometric filters that has a cosine-shaped transfer function is a delay interferometer (DI), which has been introduced in Section 2.1.2.4. Around the notch, the transfer function can be approximated as a square function, resulting to

$$T_{\text{notch}}(\nu) \sim \cos(2\pi\nu\Delta T) \sim (2\pi\nu\Delta T)^2 \tag{5.1}$$

where ν is the signal frequency and ΔT is a delay time. Based on the equation above, the optical DI performs an optical squaring function for the FM-translated duobinary signal around the notch. By allocating the 0-level of the FM duobinary bit sequence in the notch, the rectification of the bipolar duobinary signal can be performed as the -1 and $+1$ levels overlap on the quadrature point of DI transmission (Fig. 5.2). Eventually, the original 2-level on-off-keying (OOK) signal can be retrieved.

5.3.2 PAM4 Decoding

Another commonly adopted modulation format is a pulse amplitude modulation with four levels (PAM4). The four levels are represented by the bit combinations 00, 01, 10, and 11, as shown in Fig. 5.3 [159]. The PAM4 signal waveform originates from the least significant bit (LSB) signal, which is then divided into half. Afterwards, the $\frac{1}{2}$LSB is added to the most significant bit (MSB) signal. The result is four signal levels instead of two, with each signal level corresponding to a two-bit symbol. Due to that, PAM4

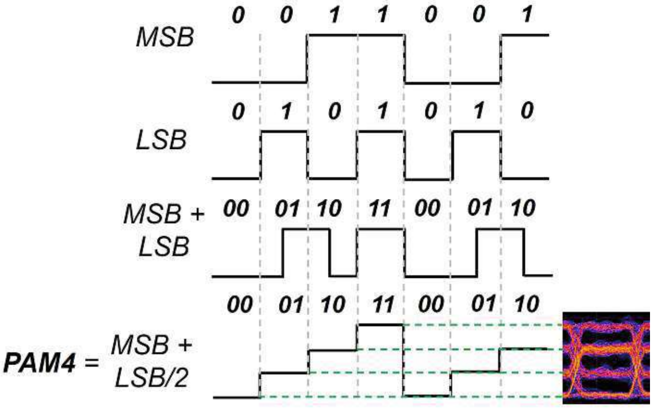

Figure 5.3: Illustration of four levels in PAM4, which is originated from a least significant bit (LSB) and a most significant bit (MSB) [159].

only requires half of the bandwidth of a given data rate by mapping every two bits into one symbol. Therefore, it supports high-speed transmission data.

In the conventional decoder, the MSB of the PAM4 signal can be recovered with one comparator in the electrical circuit with a zero threshold, while the LSB is decoded by two comparators by setting one voltage threshold to logical 1 and the other to logical 0, followed by an logic AND gate [160]-[161]. However, comparators can quickly contribute to a large power consumption [160].

In the optical domain, the decoding of the PAM4 signal can be done through a multi-level symbol slicer in the frequency domain. The idea for the decoding procedure is similar to the photonic duobinary decoding method presented above, using the interferometric filter. However, the four levels will be divided into a pair of two-bit symbols, which are LSB and MSB. Therefore, two interferometric filters are employed in this case.

Two interferometric filters with binary-staked FSR are used to slice the four levels of the PAM4 signal, as illustrated in Fig. 5.4. It is important to mention that a larger frequency deviation ν_Δ is required to cover multiple FSRs, compared to the duobinary decoding case. Therefore, the output of two filters can be directly assigned to the LSB and MSB symbols.

5.4 Concept of Neuromorphic Processing

Two neural sub-circuits are performed over the same ΦPU to demonstrate a functional photonic neuron. The neurons in the NN configuration under investigation are the same as considered in Section 4.2.2 (nodes *A*, *B*, and *C* in layer 1 as inputs and output *D* in layer 2; Fig. 4.2). Figure 5.5 shows the concept of the neuromorphic processing at

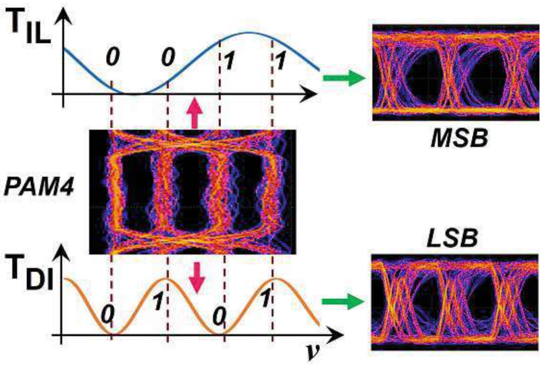

Figure 5.4: Concept of PAM4 decoding in the ΦPU using multi-level symbol slicers of interferometric filters.

the ΦPU. In the first operation, three DBR lasers are used in the linear part to send the synaptic neural signals $s_i(t)$. The magnitude of the weight $|w_i|$ is optically defined through the drive of the DBR laser, where the maximum drive current of each DBR laser is multiplied by the corresponding $|w_{\mathrm{DNN}}|$. Here, the $|w_i|$ corresponds to the $\Delta\nu$ of the peak-to-peak frequency deviation for translating the signal $s(t)$ to frequency domain $s_\nu(\nu)$ (Eq. 3.3). The polarity of weight $sgn(w_i)$ is tuned during frequency-to-intensity translation at the positive or negative interferometric slope of DI (Chapter 2). Afterwards, the summation of weighted neurons is accomplished through the wideband APD+TIA photoreceiver.

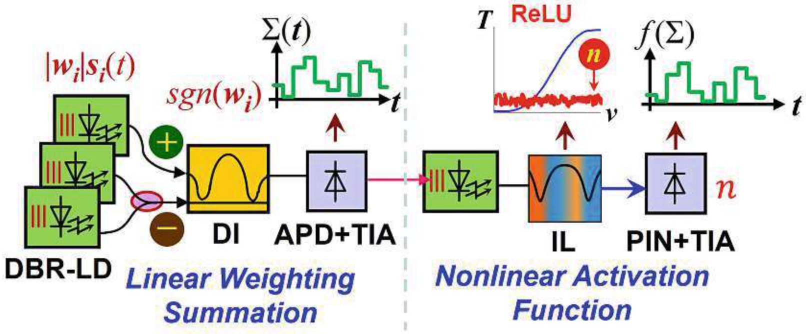

Figure 5.5: Concept of neuromorphic processing in the ΦPU consisting of two neural sub-circuits.

For the second operation, the photocurrent of the weighted sum is translated to the FM signal by directly driving the DBR laser in the nonlinear activation sub-circuit. The ReLU activation unit is achieved along the linear transfer characteristic around the quadrature point of interferometric filter IL, with the amplitude clipping due to the sensitivity n of the photodetector PIN (Chapter 3).

The following section will discuss the experimental demonstration of each signal processing task in the ΦPU architecture.

5.5 Demonstration of Photonic Processing Unit Based on Space Switch Architecture

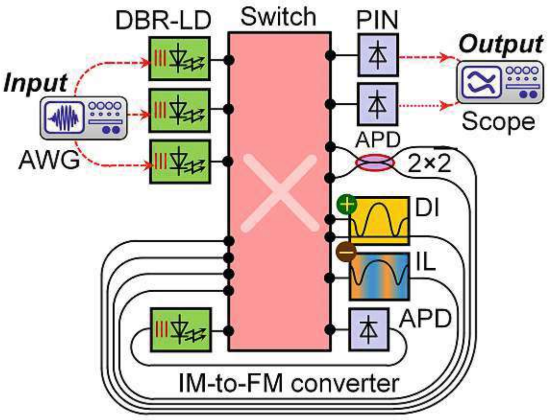

Figure 5.6: General schematic of ΦPU, with all components can be interconnected within the space switch for optical signal decoding and neuromorphic processing demonstration [MVS7].

Figure 5.6 shows the general experimental layout of the ΦPU for modulation format decoding and neuromorphic processing. The ΦPU has an 8×8 switch matrix, four DBR lasers operating at arbitrary wavelengths independent of the received optical signal, three photodetectors (PIN and APD receivers), a 2×2 splitter for signal distribution and combination, two interferometric filters, which include 10-GHz DI and 25/50 GHz IL.

The switch matrix is from DiCon MEMS 3D matrix module which is compatible for operation wavelength of 1310 nm and 1550 nm. It allows any input to be connected to any output without blocking. The insertion loss histogram of the switch module is presented in Fig. 5.7. The figure shows a considerably moderate insertion loss in the range between 0.40 and 0.63 for all input-output combinations, which is allowing an efficient all-optical cross-connect configuration.

The efficient translation of the DBR laser has been discussed in the previous chapter, and the lasers are used in the ΦPU as FM transmitters. The tuning characteristics of the

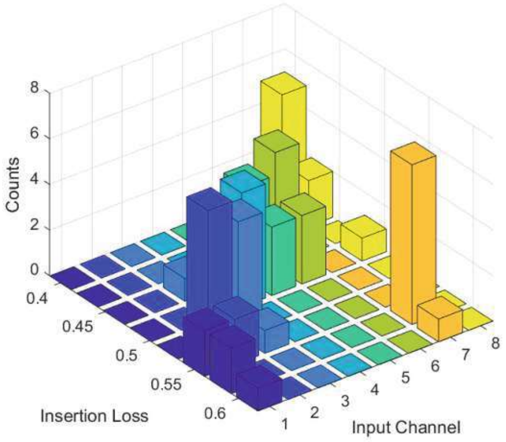

Figure 5.7: Insertion loss histogram of an 8×8 switch matrix module

four DBR lasers are presented in Fig. 5.8, following Eq. 4.4. For the neuromorphic processing, three lasers are used for weighted summation and one laser is dedicated to the activation function. The characteristic tuning parameters λ_S and I_τ for the peak-to-peak deviation in emission wavelength and concerning the FM efficiency for the DBR bias current, respectively, are in a range of $6.7 \leq \lambda_S \leq 7.5$ nm and $5.8 \leq I_\tau \leq 7.2$ mA. The slope of −7.35 GHz/mA is achievable around 1542 nm, allowing spectral processing from the electrical amplitude to the optical frequency space while ensuring a low DBR drive.

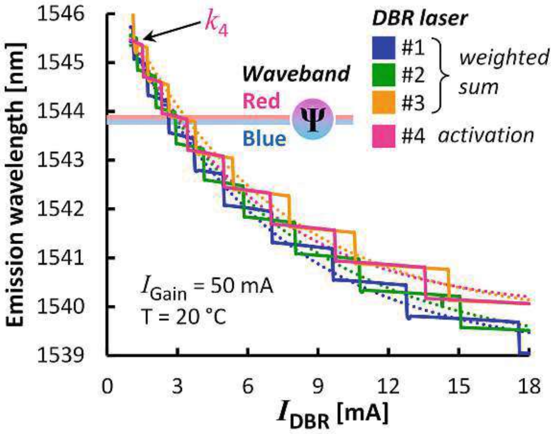

Figure 5.8: Characterisation of DBR lasers in ΦPU [MVS8].

It is important to mention that the TOSA-packaged DBR lasers have an electro-optic 3-dB bandwidth of 270 MHz for its DBR section, which is presently the main limitation since it restricts the data symbol rate R_{sym}, as mentioned in Chapter 4. Therefore, similar to the demonstration in Section 4.3.1, $R_{\mathrm{sym}} = 100\,\mathrm{Mbaud}$ is employed for the demonstration of the ΦPU functions.

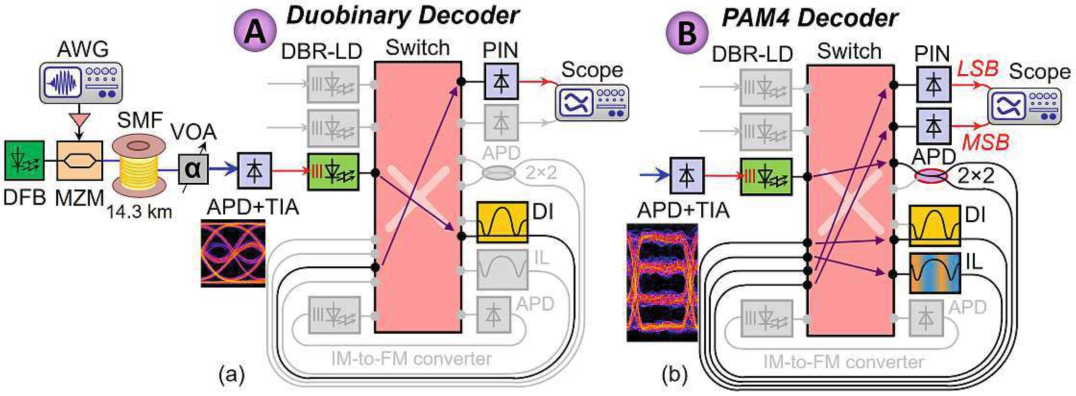

Figure 5.9: Experimental setup and ΦPU setting for (A) duobinary and (B) PAM4 decoding [MVS7].

The components in the space switch can be reconfigured for various functions. The following sub-sections presents each configuration and demonstration for a dedicated function, and eventually, their results are discussed.

5.5.1 Duobinary Reception (A)

Figure 5.9 shows the setting of ΦPU for decoding of modulation formats duobinary (A) and PAM4 (B). Generally, the ΦPU is directly fed by the optical receiver of the optical transmission link, which is an APD+TIA. In this case, it is an intensity modulation with direct detection (IM/DD) either duobinary or PAM4 transmission. The transmitted data signal is generated by an AWG that is further employed for symbol shaping and coding, and intensity-modulated on an optical carrier at 1550 nm by a MZM (Fig. 5.9(a)). The data is transmitted over 14.3 km of single-mode fiber (SMF), followed by a VOA to set the ROP.

The photocurrent of the received duobinary signal drives the DBR laser (Fig. 5.9). The frequency-translated duobinary signal is subsequently processed by the DI, which performs the optical squaring function to decode the duobinary signal in the frequency domain. Afterwards, the rectified signal is detected by a PIN photodetector and digitized through an oscilloscope to evaluate the original 2-level OOK.

Figure 5.10(a) shows the frequency-domain representation of the duobinary signal at the DBR laser output. After heterodyning with a reference laser, the three-level $(-1, 0, +1)$ duobinary pattern is clearly discernible from its heterodyned RF spectrum,

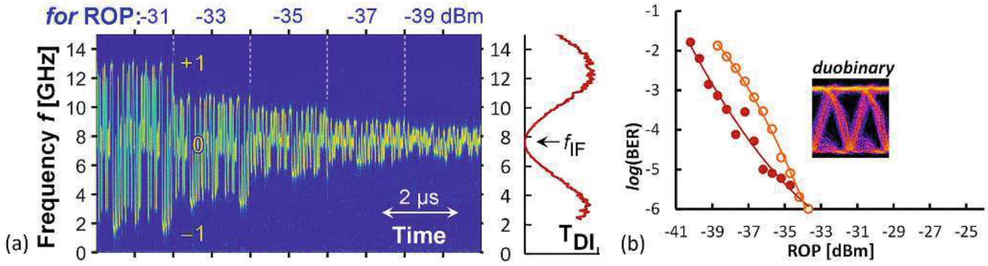

Figure 5.10: Results of duobinary reception evaluated in the ΦPU: (a) frequency-coded duobinary DBR output and (b) its BER during reception [MVS7].

where the 0-level is located at an intermediate frequency (IF) of $f_{\mathrm{IF}} = 7.6$ GHz. This IF intersects with the notch of DI, which resembles the optical duobinary decoder.

The frequency swing ν_Δ spans from the −1 to the +1 level of the duobinary signal. It can be seen from Fig. 5.10(a) that ν_Δ is getting smaller as the ROP into the photodetector decreases, meaning that the RF drive of the DBR laser (fed by APD+TIA) is reduced. It is necessary to find an optimal ν_Δ to ensure the duobinary rectification through the DI occurs between its quadrature points. To find an optimal ν_Δ, which is not too low or too high, the RF drive of the DBR laser is required to be adjusted. Figure 5.10(a) shows that the adjustment is optimal at a ROP −34 dBm for a fixed-gain TIA (Z_{T} = 2.5kΩ), corresponding to an optimal ν_Δ of ~5 GHz or equivalent to a DBR drive of about 300 µApp.

Next, the BER performance of the photonic duobinary decoding (•) is depicted in Fig. 5.10(b), including the digitally performed (DSP) duobinary decoding (○) for comparison. The reception sensitivity is −33.7 dBm at a BER of 10^{-6} when using ΦPU architecture (•). Meanwhile, the DSP-based decoding (○) does not lead to any improvement, providing that ΦPU is penalty-free.

5.5.2 PAM4 Reception (B)

Figure 5.9(b) presents the setting of the ΦPU for the application as PAM4 decoder. Similar to the case of duobinary decoding, the PAM4 signal is transmitted and received by the APD+TIA. The subsequent DBR laser translates the detected PAM4 signal to the frequency domain. Afterwards, the signal is sent to the colorless splitter for distributing the frequency-translated PAM4 signal to the multi-level slicers, consisting of DI and IL. The output of the interferometric filters are directly detected by PIN and APD at receivers, which will be digitized through the oscilloscope as LSB and MSB, respectively.

Figure 5.11(a) shows the optical spectrum where the ΦPU distributes the frequency-translated four-level signal to the 10G DI and the 25G IL using the splitter, including

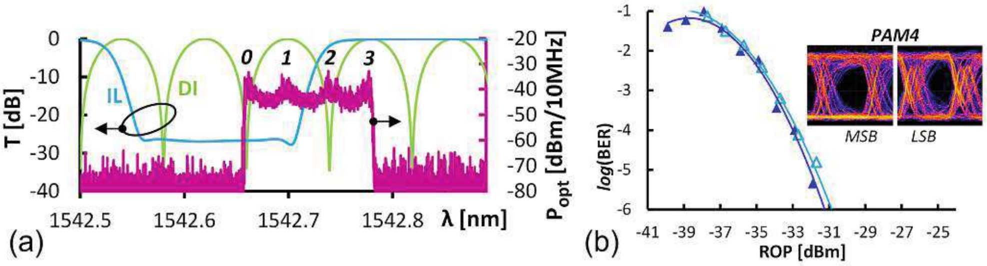

Figure 5.11: Results of photonic PAM4 decoding reception evaluated in the ΦPU: (a) PAM4 slicing and (b) its BER reception [MVS7].

the optical transmission of the filters. It can be seen that the DI and IL transfer functions build on a binary-stacked FSR. The binary filter set slices the 4-level waveform of the PAM4 signal with its symbol value [0,1,2,3] in the frequency domain.

Additionally, Fig. 5.11(b) presents the BER for the decoded MSB and LSB bit streams, showing a sensitivity of $-31.2\,\text{dBm}$ at a BER of 10^{-6} (▲). In comparison, the DSP-based multi-level symbol slicing at the APD+TIA output (△) does not show an improvement in terms of BER. Therefore, it can be confirmed again that no penalty is incurred by the ΦPU architecture.

It shall be noted that due to the smaller FSR of the 10G DI and the various transitions between the four PAM levels, the decoded LSB signal is subject to some ringing at the symbol edges as the instantaneous PAM4 frequency can swing over 1.5 FRSs (Fig. 5.11(a)). This results in a narrower eye diagram, as depicted in the inset of Fig. 5.11(b).

5.5.3 Neuromorphic Processing (C)

Figure 5.12 shows the experimental setup to evaluate the neuromorphic processing within the ΦPU. Two operations over the same ΦPU are collapsed to synthesize a neuron with three inputs in the optical domain: one dedicated to a weighted summation (solid arrows in Fig. 5.12) and another for an activation operation (dashed arrows).

The neural inputs $s_i(t)$ are sent to AWG at 100 Mbaud to drive the DBR lasers with an ultra-low drive of 130 µApp for $|w_i| = 1$. It is further translated to the frequency domain under deviation ν_Δ denoting the magnitude weight $|w_i|$. The inputs with identical polarity of weights (–) are combined via a passive combiner. Afterwards, positive and negative weights are obtained through a frequency-to-intensity conversion of the DI linear transmission slope. The weighted sum $\Sigma w_i s_i(t)$ is received by the wideband APD+TIA photodetector. This detector is part of the APD-DBR tandem serving as an intensity-to-frequency converter for the summed input, which is subsequently processed in the activation circuit. The photonic ReLU is performed in the frequency domain by

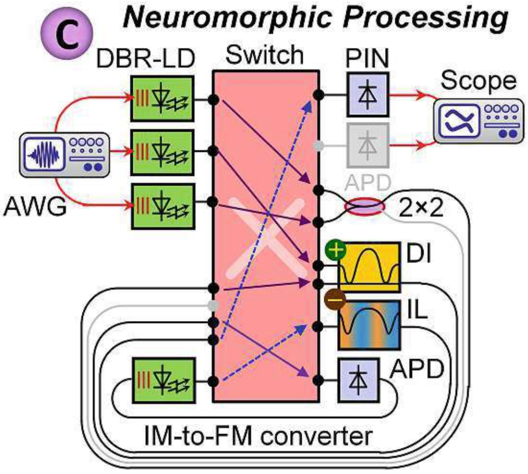

Figure 5.12: Experimental setup and ΦPU setting for neuromorphic processing function [MVS7].

using the optical IL, and the output is detected by the PIN+TIA photodetector, which additionally performs signal clipping for the rectification through the detector noise background. Finally, the output of neuromorphic processing is digitized through a real-time oscilloscope.

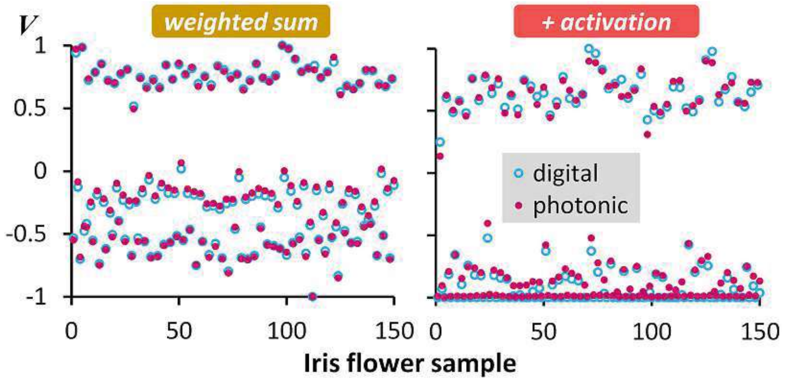

Figure 5.13: Result of neuromorphic processing evaluated in the ΦPU: normalized optical output signal compared to all-digital signal as performance benchmark [MVS8].

Figure 5.13 reports the normalized optical signal output for 150 Iris flower samples denoted by magnitude V by performing optically (•) the weighted sum and with additional the activation unit. They have an excellent agreement with the digital processing (∘), as the overlap between the optical and digital signal output is noticeable.

Additionally, Fig. 5.14(a) shows the neural signal after neuromorphic processing by computing the digital linear weighted sum and performing just the ReLU function op-

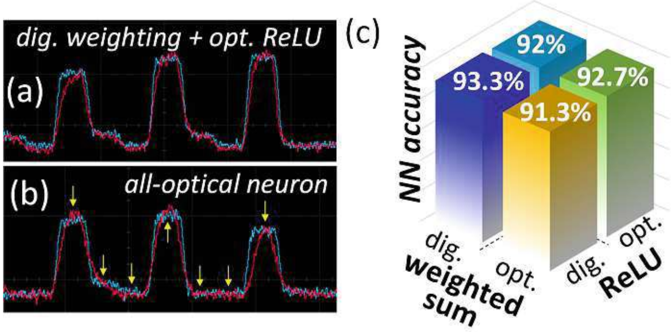

Figure 5.14: Results of neuromorphic processing evaluated in the ΦPU: neural signal comparison between optical and digital implementations, including the accuracy of Iris flower classification for multiple scenarios [162].

tically. The result shows that the optical ReLU output (red) is in good agreement with the purely digital implementation (blue). When performing both NN operations optically within ΦPU, the bandwidth limitation becomes distinguishable due to cascaded DBR lasers (Fig. 5.14(b)). However, even small pedestals in neural data are still distinguishable at the optical decision sampling points indicated by the arrows. The optical signal output of switch matrix ΦPU is fed back to NN architecture and is forwardly propagated to the output layer of the DNN to determine accuracy for the Iris flower classification. Figure 5.14(c) reports the accuracy by implementing multiple cases of neurons. The accuracies for an optical weighted sum + digital ReLU (91.3%), digital weighted sum + optical ReLU (92%), and the all-optical neuron (92.7%) are found very close to that of the all-digital neural network (93.3%). The accuracy performance in the space-switched-based ΦPU provides a benchmark for a further simplification of the neural interconnect.

The following section will describe the demonstration of the direct-drive scheme in a functional photonic neuron by employing a different neural interconnect architecture and its results.

5.6 Demonstration of Wavelength-Routed Interconnect

Other than the space-switched architecture, neuromorphic processing can also be explored in other architecture based on the wavelength-routed scheme. A wavelength-routing arrayed waveguide gratings (AWGs) is considered as a building block in this demonstration. The AWGs were first proposed by Smit [163] and are commonly used in WDM networks to multiplex/demultiplex signals. Figure 5.15 shows the schematic of an AWG. It comprises arrays of waveguides, apertures (input and output), transmitter and receiver waveguides. They can be used as wavelength routers [164], add-drop multiplexers [165], and optical cross-connect [41].

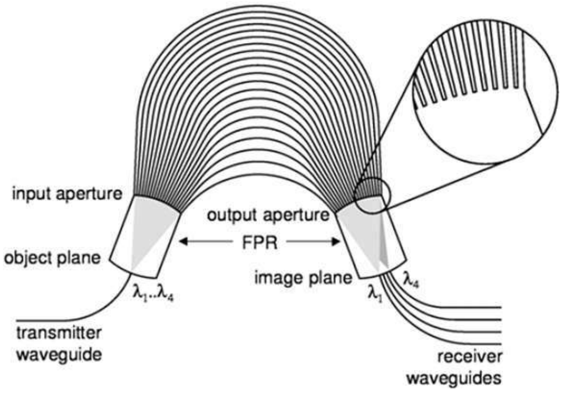

Figure 5.15: Schematic of an arrayed waveguide grating (AWG). Figure is reprinted from [166].

The working principle of AWG is based on the imaging and dispersive properties [166]. The light beam propagates to the input aperture after the input beams enter via a transmitter waveguide. It should be noted that the light in the aperture is not confined anymore; thus, it is divergent. The beams are coupled into the waveguide array and propagate towards the output aperture. The length of the array is designed where the optical path length difference between adjacent waveguides equals an integer multiple of center wavelength [166]. All the fields from the waveguides arrive at the output aperture with equal amplitude phase. Therefore, this convergence output field is transformed from the divergent beam at the input aperture. It is essential to mention that the separation to different wavelengths using AWG can be obtained by gradually increasing the lengths of the array waveguides and tilting the outgoing beam [166].

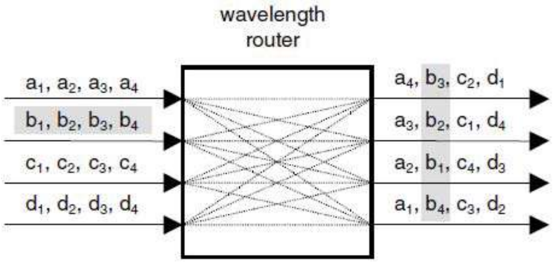

Figure 5.16: Schematic of a wavelength routing AWG. Figure is reprinted from [166].

For the wavelength routing purpose, the AWG can be designed with $N \times N$ waveguide channels, with FSR equals to $N\times$ channel spacing. Therefore, the device operates in a cyclical manner. The signal from input 1 can be designed to appear in output 1 or N, depending on the setting of frequency. For instance, a signal that vanishes from output N can reappear at output 1 if the frequency is added by an amount equal to the channel spacing [166]. Furthermore, each channel in the input AWG ports can inherently carry a number of different frequencies.

Figure 5.16 shows the illustration of the working principle of 4×4 wavelength-routed AWG. Four frequencies of signals are fed to input channel 1, denoted by signals $a_1 - a_4$ and the signal are distributed within output channels, such that output channel 4 carries the first frequency of signal a_1 and channel 1 carries the fourth frequency of signal a_4. The four frequencies from input 2 (signals $b_1 - b_4$) are distributed similarly among the outputs, but shifted in cyclic by one channel, as shown in Fig. 5.16. This cyclic behavior is the key property for multi-wavelength add-drop multiplexers [166].

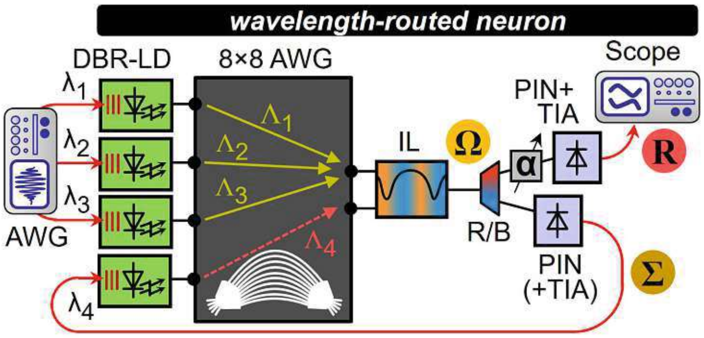

Figure 5.17: Experimental setup to evaluate a photonic neuron in the wavelength-routed architecture [MVS8].

Figure 5.17 shows the experimental setup using a cyclic 8×8 AWG with a channel spacing of 100 GHz channel according to the ITU-T G.694.1 Dense-WDM grid and a FSR of 800 GHz. The cyclic waveband behaviour for input port 1 of the AWG is depicted in Fig. 5.18. Given the eight outputs with a periodic transmission, the AWG can support many wavelengths. Furthermore, the intrinsic tunability of the DBR lasers can support the scaling of the ONN. Recent research has demonstrated an integrated AWG on a compact footprint [167]. This renders the flexibility to synthesize neural circuits by sharing the synaptic interconnect. Furthermore, a comparison will be made between employing a TIA and omitting the TIA within the tandem of the PIN+DBR laser for translating the weighted sum into the optical frequency domain. The multichannel operation is supported via a periodic filter and a wideband photodetector, which is the demarcation point within the neuron.

Two operations over the same neural interconnect architecture are collapsed to synthesize a neuron with three inputs in the optical domain: the first is dedicated to the weighted summation (solid arrows in Fig. 5.17) and the second is for the activation operation (dashed arrow), as in the space-switch architecture. Routing within the AWG is determined by the wavelength choice for several FM transmitters. The neurons in the NN configuration under investigation are the same as considered in the space-switch architecture (Fig. 4.2).

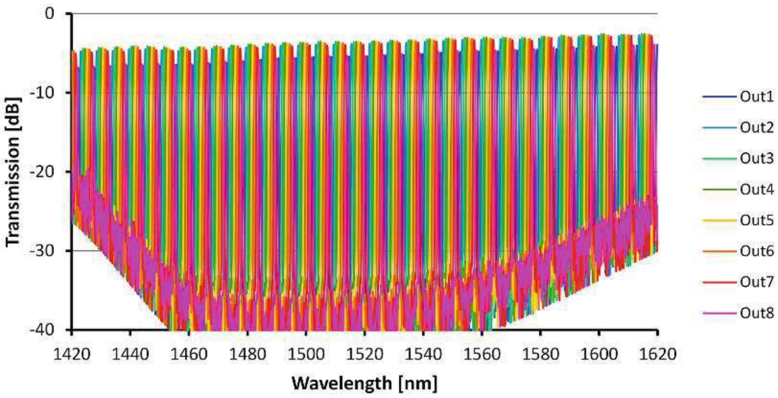

Figure 5.18: The wideband behaviour and cyclic operation of the 8×8 arrayed waveguide grating with incoming signal from the input port 1.

The signal data $s_i(t)$ (without being weighted digitally) is sent to the corresponding DBR lasers (λ_i; Fig. 5.17) at a 100 Mbaud signal rate. Three DBR lasers for linear weighted sum (solid arrows) are allocated between 1539 nm and 1543 nm. The neural inputs are frequency-modulated by the DBR laser under frequency deviation ν_Δ, reflecting the magnitude $|w_i|$ of the weight.

The weighted sum (Σ) is then received by the PIN(+TIA) receiver, which transfers the Σ according to the direct-drive scheme to the DBR laser at $\lambda_4 = 1545.7$ nm. The frequency-modulated of the Σ feeds the rectification circuit hosting the same IL (dashed arrow). Afterwards, the output ReLU is received by the second PIN+TIA photoreceiver (R). This photoreceiver is separated from the linear neural sub-circuit through use of a red/blue wavelength allocation scheme. Figure 5.17 shows a R/B (red/blue) splitter, which is utilized to separate the linear weighted sum and the nonlinear function subcircuit in the architecture, taking advantage of its complementary odd/even inputs. This will ensure no interference between the two operations. Therefore, the common IL output (Ω) is demultiplexed with the R/B splitter as the weighted sum (Σ) detected in PIN(+TIA) photodiode in the blue waveband is re-translated via the translucent concatenation approach in the optical frequency domain and the output ReLU in the red waveband is detected (R) via scope afterwards.

Figure 5.19 highlights the corresponding red/blue waveband split (Ψ) in the C-band, which is centered at 1543.8 nm. Principally, the wavelength tunable neural inputs λ_i are routed to the corresponding AWG outputs Λ_i, which are dedicated to either the linear weighted summation (blue) or the ReLU function (red). For the purpose of cost-optimization, the IL is used as the shared optical hardware (Fig. 5.19) to assign the

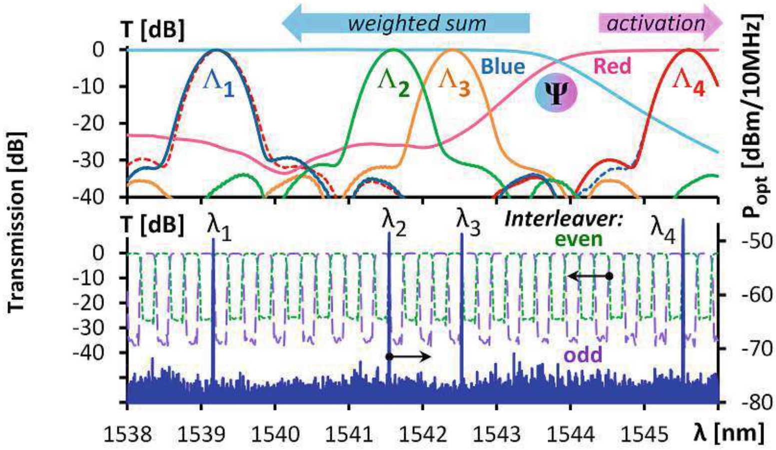

Figure 5.19: Optical spectra and relative transmission of the AWG channels for the wavelength-routed ONN at point Ω (Fig. 5.17) and the spectral allocation of DBR lasers to the IL transmission for assigning the polarity of weights [MVS8].

polarity of the weights $sgn(w_i)$ through its interferometric slope in the blue waveband and to perform ReLU operation (Section 4.2.1) in the red waveband. The neural signals are weighted with different polarity, involving one positive and two negative weights (Fig. 4.2). Figure 5.19 depicts the spectral allocation of DBR lasers to the IL transmission (even and odd channels) for weight polarity assignment and ReLU activation function.

The assignment of blue and red wavebands is motivated by the translucent concatenations scheme, since the photocurrent that drives the DBR laser (λ_4) is limited by the ROP and the responsivity (0.88 A/W) of the PIN photodiode. Therefore, the large slope efficiency k_4 = -24.7 GHz/mA at low I_{DBR} bias is preferred, as presented in Fig. 5.8.

In this demonstration, two implementations are employed for the direct-drive DBR translation of an electrical time-domain signal into an optical frequency-domain representation. A PIN+TIA photodetector is used in the first scheme after the summation (Σ, in Fig. 5.17). This means that an electrical amplification via TIA (Z_T = 2.5 kΩ) is employed when translating the detected weighted sum photocurrent into the optical frequency-coded representation. The second scheme is the direct-drive approach, where the TIA is omitted for the sake of simplicity when translating the intermediate result Σ back to the optical domain.

Figure 5.20 shows the neural signal comparison between the optical and digital subcircuit implementations. The first scenario (a) considers only performing the weighted sum in the optical domain (Σ). The neural optical output (yellow) fits well with the all-digital NN result (blue), with an average error $\overline{\varepsilon}$ of 0.09. The same good result is accomplished in (b) when comparing the photonic ReLU activation only (red, while the weighted sum is performed digitally) with the digital ReLU ($\overline{\varepsilon} = 0.20$). Additionally, the

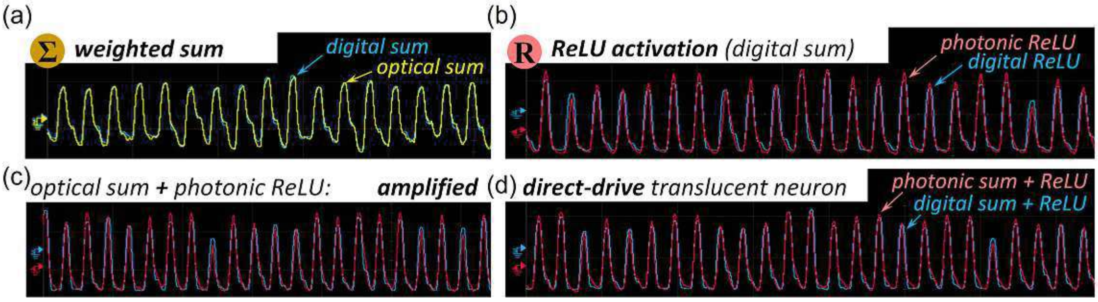

Figure 5.20: The results of neural signal comparison for optical and digital subcircuit implementations in wavelength-routed neural architecture with different scenarios [MVS8].

optical weighted sum and photonic ReLU are performed (c), where the o/e/o process in between is amplified by the TIA. The optical output signal resembles the all-digital NN with $\overline{\varepsilon} = 0.18$. Eventually, the result of the optical output signal with the direct-drive scheme (d) in the photonic neuron is in a good agreement with the all-digital NN ($\overline{\varepsilon} = 0.18$).

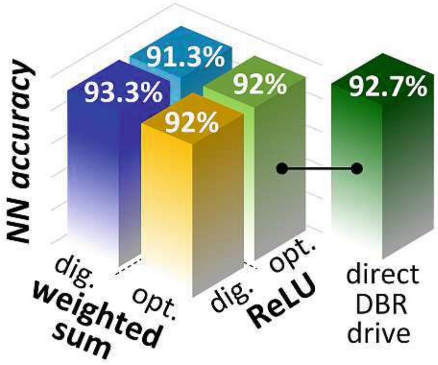

Figure 5.21: The accuracy results in wavelength-routed neural architecture with different scenarios [MVS8].

Figure 5.21 reports the accuracy of Iris flower determination for the different scenarios. The results for the optically weighed sum + digital ReLU (92%), digitally weighted sum + photonic ReLU (91.3%), and fully photonic neuron (92%) with PIN+TIA concatenation scheme are very close to the all-digital NN (93.3%). Additionally, using the direct scheme in the complete photonic neuron reaches an accuracy of 92.7%, which is a 0.6% penalty compared to the DNN. These results strongly validate of the efficiency and feasibility of using translucent concatenation scheme in the ONN. A low error can be achieved using the efficient DBR laser; therefore, the RF amplification can be omitted.

5.7 Summary

A robust photonic processor based on space switch has been demonstrated performing signal decoding and neuromorphic functions through signal translation in the frequency domain for subsequent (multi-layer) spectral processing. This has been reported in [MVS7]. The photonic circuit has been shown to operate without penalty. Furthermore, a penalty $< 2\%$ in terms of Iris flower classification accuracy has been achieved for multiple optical/digital combination of neural sub-circuits.

Additionally, the neuromorphic function in the space-switched architecture is compared with the wavelength-routed scheme based on an AWG. The cyclic and wideband properties of AWG is advantageous for performing a flexible neural interconnect. Furthermore, the good conversion efficiency of the inputs to the spectral domain permits an omission of RF electronics. The results has been published in [MVS8]. It has been demonstrated that a penalty $< 2\%$ in terms of accuracy can be accomplished too with this scheme. Especially, a minor penalty of 0.6% can be realized by performing a translucent concatenation (without TIA) between the neural sub-circuits.

An extension in electro-optic bandwidth for frequency modulation through optimized DBR lasers and an expansion of the application space is left for future work. Particularly, the AWG-based switch can be extended for other signal processing application. Some demonstrations have been shown that the space-switched and AWG-based architectures can be integrated on chip [167], [168], [169]. This allowing the possibility to scale the photonic processing in a smaller device footprint.

Chapter 6

Conclusions

Neuromorphic photonics is a burgeoning field, which holding the key to overcome the current limitations of digital electronic computing architecture, as introduced in Chapter 1. By marrying brain-inspired neural networks with photonics, the analog MAC operations can be executed at fast GHz signal rates with low latency, leveraging a massive parallelism and a vast bandwidth of photonics. This thesis makes a unique contribution to this field by investigating the potential of performing photonic neurons in the frequency domain and proposing an efficient ONN for future research directions.

This thesis has investigated the possibility of finding a set of operations in the photonics domain that are equivalent to biological neural networks, where the architecture is represented by multiple inputs of neurons that are being weighted and linearly summed and then experiencing a nonlinear activation operation for a decision boundary. The work presented here takes part in developing the weighting via frequency demodulation, as presented in Chapter 2. The first step is to realize a synaptic receptor to perform MAC through a monolithic integrated semiconductor optical amplifier (SOA) and reflective electroabsorption modulator (REAM). The periodic transfer function of the SOA-REAM device supports a wide optical bandwidth that can carry multiple wavelength channels during this frequency-to-intensity conversion process. Hence, this gives the proposed synaptic receptor an advantage for being a “colorless” demodulator. The weighting process is carried out by spectrally aligning the FM signal to the comb transfer function of the synaptic receptor. By suppressing either the mark- or space-sideband of the FM signal, a negative or positive response is accomplished for the detected signal.

The received signals at the detector, with all the weights defined during the FM demodulation of the particular wavelength channels, will be straightforwardly summed by virtue of the wide optical bandwidth of the photodetector. WDM operation has been proven through the joint demodulation and detection of two spike trains with a spectral detuning of 23 nm, thus enabling a highly parallel synaptic interconnect. Moreover, the integration of detection functionality with the weight-tunable FM demodulator has been experimentally demonstrated and confirmed through a low BER for weighted signal acquisition. It has been shown that the proposed synaptic receptor based on SOA-REAM offers an integrated solution to perform MAC functionality. A comparison is made with alternative frequency-demodulating receivers based on MRR or OTF, which reveals the possible options of FM demodulators due to a significant contrast in their transfer functions. This suggests a potential improvement in the demodulation efficiency of the SOA-REAM receptor. Further improvement can be made by integrating the EAM photodetector with an optimized amplifier such as TIA to boost the sensitivity of the synaptic receiver.

After the analog linear weighted sum is investigated, the photonic activation function has been proposed in Chapter 3. The study focuses mainly on the case of the ReLU function. The photonic ReLU is implemented through the frequency coding of neural signals using a combination of a CML laser and an interferometric IL, which is subsequently followed by a PIN photodetector. This can be done by taking advantage of the linear transmission slope of IL that mimics ReLU, supported by the natural noise floor provided by the PIN. Additionally, a comparison is demonstrated by replacing the CML+IL with the electro-optic approach of DML. Based on the evaluation, the DML approach is incompatible to perform the desired ReLU function. The photonic ReLU using CML+IL introduces an error penalty of 3% compared to DNN, which is achievable in terms of accuracy when transferring the Iris flower classification case from the digital to the optical domain. This renders the proposed CML+IL tandem suitable for a functionally complete optical neuron.

After the recipes of linear weighted summation and ReLU activation function in the photonics realm have been explored, the photonic neuron is reported based on the frequency coding scheme in Chapter 4. Multiple neural sub-circuits are collapsed over the proposed photonic ReLU function. The demonstration shows that it is possible to integrate the weighting functionality with the photonic ReLU, which can simplify the ONN implementation on a larger scale. A low penalty of 1% is achieved for this demonstration compared to the DNN. Afterwards, another scheme of the photonic neuron is demonstrated by scaling simultaneous weighting and the ReLU operation. The result shows a moderate error of 2% in accuracy, which explains the need for separation between linear and nonlinear functions.

Therefore, the efficient translucent interface scheme between the linear and nonlinear circuits in the photonic neuron is introduced and demonstrated. This can be done by employing an efficient DBR laser with a large spectral expansion to transform the signal from the amplitude to the frequency domain. This would be sufficient for the photocurrent of a preceding photodetector to drive the efficient DBR laser without the need for electrical amplification. By directly driving the DBR laser in the ReLU circuit, the accuracy of 92% for Iris flower classification is achieved. The 1% penalty compared to DNN proves that the efficient e/o/e between neural sub-circuits can be realized without a particular electrical signal conditioning.

In the last part of the thesis, Chapter 5 introduces a multi-functional photonic processing unit (ΦPU) which is implemented through a space-switched architecture. The ΦPU is penalty-free to perform signal processing such as duobinary and PAM4 decoding. Furthermore, the neuromorphic processing based on the space-switch is compared with the wavelength-routed neural interconnect using arrayed waveguide grating (AWG). The result shows that <2% of accuracy penalty is achieved for multiple combinations of optical/digital neural sub-circuits. Furthermore, a minor 0.6% penalty can be achieved as the direct drive scheme without TIA between neural sub-circuits is performed. The result supports the possibility of reducing device footprint when scaling up the translucent NN. Furthermore, the AWG-based interconnect can be further explored for multiple signal processing, such as in the ΦPU.

This thesis, dedicated to advancing the field of neuromorphic photonics, demonstrates the promise of frequency-coded photonic neurons and their potential for implementation. Looking ahead, the use of a DBR laser with a larger bandwidth could enable even higher signal processing, up to GHz implementation. Furthermore, the integration of all components on the PIC, as discussed in Chapter 4 could significantly scale up the ONN using more photonic neurons, opening up exciting possibilities for future research and development.

Bibliography

[1] D. Psaltis and N. Farhat, "Optical information processing based on an associative-memory model of neural nets with thresholding and feedback," *Optics Letters*, vol. 10, no. 2, pp. 98–100, Feb. 1985. doi: 10.1364/OL.10.000098.

[2] N. Jones, "The Information Factories," *Nature*, vol. 561, no. 7722, pp. 163–166, 2018.

[3] M. Nowell. Cisco visual networking index (VNI) forecast update. Ethernet Bandwidth Assessment Ad hoc Calls. [Accessed: 2024, April]. [Online]. Available: https://www.ieee802.org/3/ad_hoc/bwa2/public/calls/19_0624/nowell_bwa_01_190624.pdf

[4] A. S. G. Andrae and T. Edler, "On Global Electricity Usage of Communication Technology: Trends to 2030," *Challenges*, vol. 6, no. 1, pp. 117–157, 2015. doi: doi.org/10.3390/challe6010117.

[5] K. Rupp. Microprocessor trend data. [Accessed: 2024, February]. [Online]. Available: https://github.com/karlrupp/microprocessor-trend-data

[6] U. Hack. What's the real story behind the explosive growth of data? [Accessed: 2024, February]. [Online]. Available: https://www.red-gate.com/blog/database-development/whats-the-real-story-behind-the-explosive-growth-of-data

[7] Data centers. [Accessed: 2024, February]. [Online]. Available: https://www.datacentermap.com/datacenters/

[8] The data age is here. are you ready? San Francisco, CA, US. [Accessed: 2024, April]. [Online]. Available: https://www.splunk.com/en_us/pdfs/gated/ebooks/data-age.pdf

[9] J. D. Kendall and S. Kumar, "The building blocks of a brain-inspired computer," *Applied Physics Review*, vol. 7, no. 1, p. 011305, Jan. 2020. doi: 10.1063/1.5129306.

[10] T. N. Theis and H.-S. P. Wong, "The End of Moore's Law: A New Beginning for Information Technology," *Computing in Science & Engineering*, vol. 19, no. 2, pp. 41–50, Mar.-Apr. 2017. doi: 10.1109/MCSE.2017.29.

[11] N. R. B. Martins *et al.*, "Human Brain/Cloud Interface," *Frontiers Neuroscience*, vol. 13, no. 112, pp. 1–24, Mar. 2019. doi: 10.3389/fnins.2019.00112.

[12] J. Hasler and B. Marr, "Finding a roadmap to achieve large neuromorphic hardware systems," *Frontiers Neuroscience*, vol. 7, no. 118, pp. 1–29, Sep. 2013. doi: 10.3389/fnins.2013.00118.

[13] E. Strubell, A. Ganesh, and A. McCallum, “Energy and Policy Considerations for Deep Learning in NLP,” in *Proc. of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics, 2019, p. 3645–3650. doi: 10.18653/v1/P19-1355.

[14] F. Arute *et al.*, “Quantum supremacy using a programmable superconducting processor,” *Nature*, vol. 574, no. 7779, pp. 505–510, Oct. 2019. doi: 10.1038/s41586-019-1666-5.

[15] J. von Neumann, “First draft of a report on the EDVAC,” *Annals of the History of Computing*, vol. 15, no. 4, pp. 27–75, 1993. doi: 10.1109/85.238389.

[16] P. W. Shor, “Algorithms for quantum computation: Discrete logarithms and factoring,” in *Proc. of the 35th Annual Symposium on Foundations of Computer Science*. Santa Fe, NM, USA: IEEE, Nov. 1994, pp. 124–134. doi: 10.1109/SFCS.1994.365700.

[17] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O’Brien, “Quantum computers,” *Nature*, vol. 464, no. 7285, pp. 45–53, Mar. 2010. doi: 10.1038/nature08812.

[18] C. D. Schuman, S. R. Kulkarni, M. Parsa, J. P. Mitchell, P. Date, and B. Kay, “Opportunities for neuromorphic computing algorithms and applications,” *Nature Computational Science*, vol. 2, no. 1, pp. 10–19, Jan. 2022. doi: 10.1038/s43588-021-00184-y.

[19] T. Baba *et al.*, “Proposal of disruptive computing (A computing-domain oriented approach,” *Japanese Journal of Applied Physics*, vol. 59, no. 5, p. 050503, May 2020. doi: 10.35848/1347-4065/ab8577.

[20] C. Mead, *Analog VLSI and Neural Systems*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1989. ISBN 0201059924

[21] K. Vandoorne, J. Dambre, D. Verstraeten, B. Schrauwen, and P. Bienstman, “Parallel reservoir computing using optical amplifiers,” *IEEE Transactions on Neural Networks*, vol. 22, no. 9, pp. 1469–1481, Sep. 2011. doi: 10.1109/TNN.2011.2161771.

[22] Q. Xu and R. Soref, “Reconfigurable optical directed-logic circuits using microresonator-based optical switches,” *Optics Express*, vol. 19, no. 6, pp. 5244–5259, Mar. 2011. doi: 10.1364/OE.19.005244.

[23] P. A. Merolla *et al.*, “A million spiking-neuron integrated circuit with a scalable communication network and interface,” *Science*, vol. 345, no. 6197, p. 668–673, Aug. 2014. doi: 10.1126/science.1254642.

[24] J. Hsu, “IBM’s new brain [news],” *IEEE Spectrum*, vol. 51, no. 10, pp. 17–19, Oct. 2014. doi: 10.1109/MSPEC.2014.6905473.

[25] C. A. Thraskias *et al.*, “Survey of photonic and plasmonic interconnect technologies for intra-datacenter and high-performance computing communications,” *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2758–2783, 2018. doi: 10.1109/COMST.2018.2839672.

[26] B. V. Benjamin *et al.*, "Neurogrid: A mixed-analog-digital multichip system for large-scale neural simulations," *Proceedings of the IEEE*, vol. 102, no. 5, pp. 699–716, May 2014. doi: 10.1109/JPROC.2014.2313565.

[27] J. Schemmel, D. Brüderle, A. Grübl, M. Hock, K. Meier, and S. Millner, "A wafer-scale neuromorphic hardware system for large-scale neural modeling," *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, pp. 1947–1950, May - Jun. 2010. doi: 10.1109/ISCAS.2010.5536970.

[28] S. Furber, "Large-scale neuromorphic computing systems," *Journal of Neural Engineering*, vol. 13, no. 051001, pp. 1–14, Aug. 2016. doi: 10.1088/1741-2560/13/5/051001.

[29] Y. Shen *et al.*, "Silicon photonics for extreme scale systems," *Journal of Lightwave Technology*, vol. 37, no. 2, pp. 245–259, Jan. 2019. doi: 10.1109/JLT.2019.2897365.

[30] X. Guo, J. Xiang, Y. Zhang, and Y. Su, "Integrated neuromorphic photonics: Synapses, neurons, and neural networks," *Advanced Photonics Research*, vol. 2, no. 6, p. 2000212, Mar. 2021. doi: https://doi.org/10.1002/adpr.202000212.

[31] E. Goi, Q. Zhang, X. Chen, H. Luan, and M. Gu, "Perspective on photonic memristive neuromorphic computing," *PhotoniX*, vol. 1, no. 3, pp. 1–26, Mar. 2020. doi: 10.1186/s43074-020-0001-6.

[32] M. Nahmias, B. Shastri, A. Tait, T. De Lima, and P. Prucnal, "Neuromorphic photonics," pp. 34–41, 2018.

[33] T. De Lima *et al.*, "Machine learning with neuromorphic photonics," *Journal of Lightwave Technology*, vol. 37, no. 5, pp. 1515–1534, Mar. 2019. doi: 10.1109/JLT.2019.2903474.

[34] W. Zhang *et al.*, "Broadband physical layer cognitive radio with an integrated photonic processor for blind source separation," *Nature communications*, vol. 14, no. 1, p. 1107, 2023. doi: 10.1038/s41467-023-36814-4.

[35] Y. Bai *et al.*, "Photonic Multiplexing Techniques for Neuromorphic Computing," *Nanophotonics*, vol. 12, no. 5, pp. 795–817, Jan. 2023. doi: doi.org/10.1515/nanoph-2022-0485.

[36] Y. Shen *et al.*, "Deep learning with coherent nanophotonic circuits," *Nature Photonics*, vol. 11, pp. 441–446, Jun. 2017. doi: 10.1038/nphoton.2017.93.

[37] F. Shokraneh, S. Gagnon, M. Nezami, and O. Liboiron-Ladouceur, "A single layer neural network implemented by a 4×4 MZI-based optical processor," *IEEE Photonics Journal*, vol. 11, no. 6, pp. 1–12, Dec. 2019. doi: 10.1109/JPHOT.2019.2952562.

[38] F. Ashtiani, A. J. Geers, and F. Aflatouni, "An on-chip photonic deep neural network for image classification," *Nature*, vol. 606, no. 7914, pp. 501–506, Jun. 2022. doi: 10.1038/s41586-022-04714-0.

[39] W. Bogaerts *et al.*, "Silicon microring resonators," *Laser & Photonics Review*, vol. 6, no. 1, p. 47–73, Sep. 2012. doi: 10.1002/lpor.201100017.

[40] A. Tait, M. Nahmias, B. Shastri, and P. Prucnal, “Broadcast and weight: An integrated network for scalable photonic spike processing,” *Journal of Lightwave Technology*, vol. 32, no. 21, pp. 4029–4041, Nov. 2014. doi: 10.1109/JLT.2014.2345652.

[41] B. Shi, N. Calabretta, and R. Stabile, “Deep neural network through an InP SOA-based photonic integrated cross-connect,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 1, p. 7701111, Jan.-Feb. 2020. doi: 10.1109/JSTQE.2019.2945548.

[42] R. Hamerly, L. Bernstein, A. Sludds, M. Soljačić, and D. Englund, “Large-scale optical neural networks based on photoelectric multiplication,” *Phys. Rev. X*, vol. 9, p. 021032, May 2019. doi: 10.1103/PhysRevX.9.021032.

[43] Y. Yang, P. Zhou, P. Mu, and N. Li, “Time-delayed reservoir computing based on an optically pumped spin VCSEL for high-speed processing,” *Nonlinear Dynamics*, vol. 107, no. 3, pp. 2619–2632, Jan. 2022. doi: 10.1007/s11071-021-07140-5.

[44] M. Borghi, S. Biasi, and L. Pavesi, “Reservoir computing based on a silicon microring and time multiplexing for binary and analog operations,” *Scientific Reports*, vol. 11, no. 1, p. 15642, Aug. 2021. doi: 10.1038/s41598-021-94952-5.

[45] A. Nguyen, K. Pham, D. Ngo, T. Ngo, and L. Pham, “An analysis of state-of-the-art activation functions for supervised deep neural network,” in *Proc. International Conference System Science and Engineering (ICSSE)*, Ho Chi Minh City, Vietnam, Aug. 2021, pp. 215–220. doi: 10.1109/ICSSE52999.2021.9538437.

[46] O. Destras, S. Le Beux, F. G. De Magalhães, and G. Nicolescu, “Survey on activation functions for optical neural networks,” *ACM Comput. Surv.*, vol. 56, no. 2, Feb. 2024. doi: 10.1145/3607533.

[47] B. Schrenk, “Electroabsorption-modulated laser as optical transmitter and receiver: status and opportunities,” *IET Optoelectron.*, vol. 14, no. 6, pp. 374–385, Oct. 2020. doi: 10.1049/iet-opt.2020.0010.

[48] J. George, A. Mehrabian, R. Amin, P. R. Prucnal, T. El-Ghazawi, and V. J. Sorger, “Neural network activation functions with electro-optic absorption modulators,” in *Proc. 2018 IEEE International Conference on Rebooting Computing (ICRC)*, Nov. 2018, pp. 1–5. doi: 10.1109/ICRC.2018.8638590.

[49] J. George *et al.*, “Neuromorphic photonics with electro-absorption modulators,” *Optics Express*, vol. 27, no. 4, pp. 5181–5191, Feb. 2019. doi: 10.1364/OE.27.005181.

[50] A. Tait *et al.*, “Silicon photonic modulator neuron,” *Physical Review Applied*, vol. 11, no. 6, p. 064043, Jun. 2019. doi: 10.1103/PhysRevApplied.11.064043.

[51] I. Williamson, T. Hughes, M. Minkov, B. Bartlett, S. Pai, and S. Fan, “Reprogrammable electro-optic nonlinear activation functions for optical neural networks,” *IEEE J. Select. Top. Quant. Electr.*, vol. 26, no. 1, pp. 1–12, Jan.-Feb. 2019. doi: 10.1109/JSTQE.2019.2930455.

[52] M. M. P. Fard *et al.*, “Experimental realization of arbitrary activation functions for optical neural networks,” *Opt. Express*, vol. 28, no. 8, pp. 12 138–12 148, Apr. 2020. doi: 10.1364/OE.391473.

[53] J. R. Rausell Campo and D. Pérez-López, “Reconfigurable activation functions in integrated optical neural networks,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 28, no. 4: Mach. Learn. in Photon. Commun. and Meas. Syst., pp. 1–13, Jul.-Aug. 2022. doi: 10.1109/JSTQE.2022.3169833.

[54] Y. Huang, W. Wang, L. Qiao, X. Hu, and T. Chu, “Programmable low-threshold optical nonlinear activation functions for photonic neural networks,” *Opt. Lett.*, vol. 47, no. 7, pp. 1810–1813, Apr. 2022. doi: 10.1364/OL.451287.

[55] G. Mourgias-Alexandris, A. Tsakyridis, N. Passalis, A. Tefas, K. Vyrsokinos, and N. Pleros, “An all-optical neuron with sigmoid activation function,” *Optics Express*, vol. 27, no. 7, pp. 9620–9630, Apr. 2019. doi: 10.1364/OE.27.009620.

[56] B. Shi, N. Calabretta, and R. Stabile, “InP photonic integrated multi-layer neural networks: Architecture and performance analysis,” *APL Photonics*, vol. 7, no. 1, p. 010801, Jan. 2022. doi: 10.1063/5.0066350.

[57] Q. Li *et al.*, “SOA-based all-optical neuron with reconfigurable nonlinear activation functions,” in *Conference on Lasers and Electro-Optics*. Optica Publishing Group, 2022, p. SF4F.6.

[58] A. Jha, C. Huang, and P. Prucnal, “Reconfigurable all-optical nonlinear activation functions for neuromorphic photonics,” *Optics Letters*, vol. 45, no. 17, pp. 4819–4822, Sep. 2020. doi: 10.1364/OL.398234.

[59] Z. Fu, Z. Wang, P. Bienstman, R. Jiang, J. Wang, and C. Wu, “Programmable low-power consumption all-optical nonlinear activation functions using a micro-ring resonator with phase-change materials,” *Opt. Express*, vol. 30, no. 25, pp. 44 943–44 953, Dec. 2022. doi: 10.1364/OE.476110.

[60] B. Wu, H. Li, W. Tong, J. Dong, and X. Zhang, “Low-threshold all-optical nonlinear activation function based on a Ge/Si hybrid structure in a microring resonator,” *Opt. Mater. Express*, vol. 12, no. 3, pp. 970–980, Mar. 2022. doi: 10.1364/OME.447330.

[61] N. Takahashi *et al.*, “Optical ReLU using membrane lasers for an all-optical neural network,” *Opt. Lett.*, vol. 47, no. 21, pp. 5715–5718, Nov. 2022. doi: 10.1364/OL.471603.

[62] J. Crnjanski, M. Krstić, A. Totović, N. Pleros, and D. Gvozdić, “Adaptive sigmoid-like and PReLU activation functions for all-optical perceptron,” *Opt. Lett.*, vol. 46, no. 9, pp. 2003–2006, May. 2021. doi: 10.1364/OL.422930.

[63] G. Li *et al.*, “All-optical ultrafast ReLU function for energy-efficient nanophotonic deep learning,” *Nanophotonics*, vol. 12, no. 5, pp. 847–855, May 2022. doi: 10.1515/nanoph-2022-0137.

[64] B. Shastri, T. de Lima, A. Tait, M. Nahmias, and P. Prucnal, “Emergence of neuromorphic photonics,” in *Advanced Photonics 2017*. Optica Publishing Group (formerly OSA), Jul. 2017. doi: 10.1364/PS.2017.PTu3C.4. Paper PTu3C.4.

[65] H. van Appeven. How to build brain-inspired neural networks based on light. Eindhoven, Netherlands. [Accessed: 2024, February]. [Online]. Available: https://www.tue.nl/en/news-and-events/news-overview/12-04-2022-how-to-build-brain-inspired-neural-networks-based-on-light

[66] J. J. Martínez, J. G. Gregorio, A. L. Lucia, A. V. Velasco, J. C. Aguado, and M. Ángeles Losada Binué, "Novel WDM-PON architecture based on a spectrally efficient im-fsk scheme using DMLs and RSOAs," *J. Lightwave Technol.*, vol. 26, no. 3, pp. 350–356, Feb. 2008. doi: 10.1109/JLT.2007.909864.

[67] C. del Río Campos and P. R. Horche, *Effects of Dispersion Fiber on CWDM Directly Modulated System Performance*. Rijeka: IntechOpen, 2012, ch. 3.

[68] J. Yu *et al.*, "Applications of 40-Gb/s chirp-managed laser in access and metro networks," *Journal of Lightwave Technology*, vol. 27, no. 3, pp. 253–265, Feb. 2009. doi: 10.1109/JLT.2008.2006064.

[69] G. Yabre, "Improved direct-modulation characteristics of a semiconductor laser by fm/im conversion through an interferometer," *Journal of Lightwave Technology*, vol. 14, no. 10, pp. 2135–2140, Oct. 1996. doi: 10.1109/50.541200.

[70] S. L. Woodward, "Lightwave CATV systems using frequency-modulated laser and interferometer," *Electronics Letters*, vol. 25, no. 24, p. 1665–1666, Nov. 1989. doi: 10.1049/el:19891116.

[71] B. E. A. Saleh and M. C. Teich, *Resonator Optics*. John Wiley & Sons, Ltd, 1991, ch. 9, pp. 310–341. ISBN 9780471839651

[72] N. H. Zhu *et al.*, "Directly modulated semiconductor lasers," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 1, pp. 1–19, Jan.-Feb. 2018. doi: 10.1109/JSTQE.2017.2720959.

[73] F. Fan and D. Mahgerefteh, "Chirp managed lasers," *Optik & Photonik*, vol. 2, no. 4, pp. 39–41, Mar. 2011. doi: https://doi.org/10.1002/opph.201190284.

[74] D. Mahgerefteh, Y. Matsui, X. Zheng, and K. McCallion, "Chirp managed laser and applications," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 5, pp. 1126–1139, Sep.-Oct. 2010. doi: 10.1109/JSTQE.2009.2037336.

[75] C. W. Lee, O. K. Kwon, and K. S. Kim, "Wide-wavelength-tunable distributed bragg reflector laser diode with high thermal efficiency," *Opt. Express*, vol. 31, no. 20, pp. 32 126–32 133, Sep. 2023. doi: 10.1364/OE.497943.

[76] K. Lawniczuk, "Multiwavelength transmitters in generic photonic integration technologies," Phd Thesis (Research TU/e / Graduation TU/e), Electrical Engineering. ISBN 978-90-386-3550-7 2014. doi: 10.6100/IR763106.

[77] B. Mason, G. Fish, S. DenBaars, and L. Coldren, "Widely tunable sampled grating DBR laser with integrated electroabsorption modulator," *IEEE Photonics Technology Letters*, vol. 11, no. 6, pp. 638–640, Jun. 1999. doi: 10.1109/68.766769.

[78] L. Han *et al.*, “Electroabsorption-modulated widely tunable DBR laser transmitter for WDM-PONs,” *Opt. Express*, vol. 22, no. 24, pp. 30 368–30 376, Dec. 2014. doi: 10.1364/OE.22.030368.

[79] D. Zhao, “High-precision distributed bragg reflectors in a generic photonic integration platform,” Phd Thesis (Research TU/e / Graduation TU/e), Electrical Engineering. ISBN 978-90-386-4627-5 2018.

[80] T. Erdogan, “Fiber grating spectra,” *Journal of Lightwave Technology*, vol. 15, no. 8, pp. 1277–1294, Aug. 1997. doi: 10.1109/50.618322.

[81] N. T. Islam, C. C. W. Lim, C. Cahall, J. Kim, and D. J. Gauthier, “Provably secure and high-rate quantum key distribution with time-bin qudits,” *Science Advances*, vol. 3, no. 11, p. e1701491, Nov. 2017. doi: 10.1126/sciadv.1701491.

[82] E. Dejband *et al.*, “Utilizing a tunable delay line interferometer to improve the sensing accuracy of an FBG sensor system,” *Photonics*, vol. 9, no. 11, Nov. 2022. doi: 10.3390/photonics9110869.

[83] W. Wu, Y. Yu, S. Hu, B. Zou, and X. Zhang, “All-optical format conversion for polarization and wavelength division multiplexed system,” *IEEE Photonics Technology Letters*, vol. 24, no. 18, pp. 1606–1609, Sep. 2012. doi: 10.1109/LPT.2012.2210543.

[84] F. Bontempi *et al.*, “Multifunctional current-controlled InP photonic integrated delay interferometer,” *IEEE Journal of Quantum Electronics*, vol. 48, no. 11, pp. 1453–1461, Nov. 2012. doi: 10.1109/JQE.2012.2219038.

[85] J. Li, K. Worms, R. Maestle, D. Hillerkuss, W. Freude, and J. Leuthold, “Free-space optical delay interferometer with tunable delay and phase,” *Opt. Express*, vol. 19, no. 12, pp. 11 654–11 666, Jun. 2011. doi: 10.1364/OE.19.011654.

[86] A. Ayazi, T. Baehr-Jones, Y. Liu, A. E.-J. Lim, and M. Hochberg, “Linearity of silicon ring modulators for analog optical links,” *Opt. Express*, vol. 20, no. 12, pp. 13 115–13 122, Jun. 2012. doi: 10.1364/OE.20.013115.

[87] P. Steglich *et al.*, “Silicon photonic micro-ring resonators for chemical and biological sensing: A tutorial,” *IEEE Sensors Journal*, vol. 22, no. 11, pp. 10 089–10 105, Jun. 2022. doi: 10.1109/JSEN.2021.3119547.

[88] I. C. M. Littler, M. Rochette, and B. J. Eggleton, “Adjustable bandwidth dispersionless bandpass fbg optical filter,” *Opt. Express*, vol. 13, no. 9, pp. 3397–3407, May 2005. doi: 10.1364/OPEX.13.003397.

[89] K. Yu, W. Liu, D. Huang, and J. Chang, “C-band three-port tunable band-pass thin film optical filter with low polarization-sensitivity,” *Optics Communications*, vol. 281, no. 14, pp. 3709–3714, Jul. 2008. doi: https://doi.org/10.1016/j.optcom.2008.03.019.

[90] L. Domash, M. Wu, N. Nemchuk, and E. Ma, “Tunable and switchable multiple-cavity thin film filters,” *Journal of Lightwave Technology*, vol. 22, no. 1, pp. 126–135, Jan. 2004. doi: 10.1109/JLT.2004.823349.

[91] J.-W. Jeong, I. W. Jung, H. J. Jung, D. M. Baney, and O. Solgaard, "Multi-functional tunable optical filter using mems spatial light modulator," *Journal of Microelectromechanical Systems*, vol. 19, no. 3, pp. 610–618, Jun. 2010. doi: 10.1109/JMEMS.2010.2043641.

[92] D. Mahgerefteh, "Chirp managed laser (CML) and applications," 2011.

[93] Z. Li, L. Yi, and W. Hu, "Comparison of downstream transmitters for high loss budget of long-reach 10G-PON," in *OFC 2014*, Mar. 2014, pp. 1–3. doi: 10.1364/OFC.2014.Tu2C.4.

[94] J. Raring, E. Skogen, L. Johansson, M. Sysak, S. DenBaars, and L. Coldren, "Widely tunable negative-chirp SG-DBR laser/EA-modulated transmitter," *Journal of Lightwave Technology*, vol. 23, no. 1, pp. 80–86, Jan. 2005. doi: 10.1109/JLT.2004.840330.

[95] Delay line interferometers MINT and WT-MINT. [Accessed: 2024, February]. [Online]. Available: https://kylia.com/api-website-feature/files/download/11118/datasheet-MINT-V12.pdf

[96] Adjustable bandwidth tunable filter OTF-350. [Accessed: 2024, February]. [Online]. Available: https://www.santec.com/dcms_media/other/instruments_OTF-350-C-E.pdf

[97] C. W. F. Parsonson, Z. Shabka, W. K. Chlupka, B. Goh, and G. Zervas, "Optimal control of SOAs with artificial intelligence for sub-nanosecond optical switching," *Journal of Lightwave Technology*, vol. 38, no. 20, pp. 5563–5573, Oct. 2020. doi: 10.1109/JLT.2020.3004645.

[98] A. Sobhanan *et al.*, "Semiconductor optical amplifiers: recent advances and applications," *Adv. Opt. Photon.*, vol. 14, no. 3, pp. 571–651, Sep. 2022. doi: 10.1364/AOP.451872.

[99] M. Connelly, "Semiconductor optical amplifiers and their applications," in *presented at 3rd Spanish Meeting of Optoelectronics, OPTOEL*, vol. 3, 2003.

[100] M. O'Mahony, "Semiconductor laser optical amplifiers for use in future fiber systems," *Journal of Lightwave Technology*, vol. 6, no. 4, pp. 531–544, Apr. 1988. doi: 10.1109/50.4035.

[101] R. Welstand, S. Pappert, C. Sun, J. Zhu, Y. Liu, and P. Yu, "Dual-function electroabsorption waveguide modulator/detector for optoelectronic transceiver applications," *IEEE Photonics Technology Letters*, vol. 8, no. 11, pp. 1540–1542, Nov. 1996. doi: 10.1109/68.541576.

[102] G. L. Li, P. K. L. Yu, and I. Tutorial, "Optical intensity modulators for digital and analog applications," *J. Lightwave Technol.*, vol. 21, no. 9, pp. 2010–2030, Sep. 2003. doi: 10.1109/JLT.2003.815654.

[103] M. Trajkovic, "High speed electro-absorption modulators in indium phosphide generic integration technologies," Phd Thesis 1 (Research TU/e / Graduation TU/e), Electrical Engineering. ISBN 978-90-386-4758-6 May 2019.

[104] O. Mitomi, S. Nojima, I. Kotaka, K. Wakita, K. Kawano, and M. Naganuma, "Chirping characteristic and frequency response of mqw optical intensity modulator," *Journal of Lightwave Technology*, vol. 10, no. 1, pp. 71–77, 1992. doi: 10.1109/50.108739.

[105] R. Salvatore, R. Sahara, M. Bock, and I. Libenzon, "Electroabsorption modulated laser for long transmission spans," *IEEE Journal of Quantum Electronics*, vol. 38, no. 5, pp. 464–476, May 2002. doi: 10.1109/3.998618.

[106] H. Dalir, Y. Takahashi, and F. Koyama, "Compact low driving voltage (< 400 mVpp) electro-absorption modulator laterally integrated with VCSEL," *Electronics letters*, vol. 50, no. 2, pp. 101–103, Jan. 2014. doi: 10.1049/el.2013.3662.

[107] J. Fujikata *et al.*, "High-performance Ge/Si electro-absorption optical modulator up to 85°C and its highly efficient photodetector operation," *Opt. Express*, vol. 31, no. 6, pp. 10 732–10 743, Mar. 2023. doi: 10.1364/OE.484380.

[108] A. Garreau *et al.*, "10Gbit/s amplified reflective electroabsorption modulator for colourless access networks," in *2006 International Conference on Indium Phosphide and Related Materials Conference Proceedings*, May 2006, pp. 168–170. doi: 10.1109/ICIPRM.2006.1634139. Paper TuA2.3.

[109] J. Wang, H. Olesen, and K. Stubkjaer, "Recombination, gain and bandwidth characteristics of 1.3-μm semiconductor laser amplifiers," *Journal of Lightwave Technology*, vol. 5, no. 1, pp. 184–189, Jan. 1987. doi: 10.1109/JLT.1987.1075407.

[110] B. Schrenk *et al.*, "An all-optical carrier recovery scheme for access networks with simple ASK modulation," *Journal of Optical Communications and Networking*, vol. 3, no. 9, pp. 704–712, Sep. 2011. doi: 10.1364/JOCN.3.000704.

[111] Z. Bielecki, K. Achtenberg, M. E. Kopytko, J. A. Mikołajczyk, J. Wojtas, and A. W. Rogalski, "Review of photodetectors characterization methods," *Bulletin of the Polish Academy of Sciences, Technical Sciences*, vol. 70, no. 2, 2022. doi: 10.24425/bpasts.2022.140534.

[112] V. Mackowiak, J. Peupelmann, Y. Ma, and A. Gorges, "NEP–noise equivalent power," *Thorlabs, Inc*, vol. 56, 2015.

[113] A. Kakkar. IRIS-flower-classification. [Accessed: 2024, February]. [Online]. Available: https://github.com/amberkakkar01/IRIS-Flower-classification?tab=readme-ov-file

[114] R. Fisher, "The use of multiple measurements in taxonomic problems," *Annual of Eugenics*, vol. 7, no. 2, pp. 179–188, Sep. 1936. doi: 10.1111/j.1469-1809.1936.tb02137.x.

[115] E. Anderson, "The species problem in Iris," *Annals of the Missouri Botanical Garden*, vol. 23, no. 3, pp. 457–509, 1936. doi: 10.2307/2394164.

[116] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017.

[117] L. Pham and A. Schindler, "Jollybee project report (v0.2 300922): Multilayer perception neural network model for Iris dataset," 2022, 30 September 2022.

[118] G. P. Agrawal, “Chapter 7 - fiber-optic communications,” in *Applications of Nonlinear Fiber Optics (Third Edition)*, 3rd ed., G. P. Agrawal, Ed. Academic Press, 2021, pp. 309–368. ISBN 978-0-12-817040-3. doi: https://doi.org/10.1016/B978-0-12-817040-3.00013-4.

[119] K. Kikuchi, “Fundamentals of coherent optical fiber communications,” *Journal of lightwave technology*, vol. 34, no. 1, pp. 157–179, Jan. 2016. doi: 10.1109/JLT.2015.2463719.

[120] L. Coldren, G. Fish, Y. Akulova, J. Barton, L. Johansson, and C. Coldren, “Tunable semiconductor lasers: a tutorial,” *Journal of Lightwave Technology*, vol. 22, no. 1, pp. 193–202, Jan. 2004. doi: 10.1109/JLT.2003.822207.

[121] F. Kennis, “The tunable three section distributed bragg reflector laser,” 1996, master Thesis. [Online]. Available: https://research.tue.nl/en/studentTheses/the-tunable-three-section-distributed-bragg-reflector-laser

[122] J. Wang, F. Sciarrino, A. Laing, and M. G. Thompson, “Integrated photonic quantum technologies,” *Nature Photonics*, vol. 14, no. 5, pp. 273–284, 2020. doi: 10.1038/s41566-019-0532-1.

[123] M. Moralis-Pegios *et al.*, “Neuromorphic silicon photonics and hardware-aware deep learning for high-speed inference,” *Journal of Lightwave Technology*, vol. 40, no. 10, pp. 3243–3254, May 2022. doi: 10.1109/JLT.2022.3171831.

[124] M. Al-Qadasi, L. Chrostowski, B. Shastri, and S. Shekhar, “Scaling up silicon photonic-based accelerators: Challenges and opportunities,” *APL Photonics*, vol. 7, no. 2, Feb. 2022. doi: 10.1063/5.0070992.

[125] B. Bai, H. Shu, X. Wang, and W. Zou, “Towards silicon photonic neural networks for artificial intelligence,” *Science China Information Sciences*, vol. 63, pp. 1–14, May 2020. doi: 10.1007/s11432-020-2872-3.

[126] X. Xiao, M. B. On, T. Van Vaerenbergh, D. Liang, R. G. Beausoleil, and S. Yoo, “Large-scale and energy-efficient tensorized optical neural networks on III–V-on-silicon MOSCAP platform,” *Apl Photonics*, vol. 6, no. 12, Dec. 2021. doi: 10.1063/5.0070913.

[127] A. Jha, C. Huang, H.-T. Peng, B. Shastri, and P. R. Prucnal, “Photonic spiking neural networks and graphene-on-silicon spiking neurons,” *Journal of Lightwave Technology*, vol. 40, no. 9, pp. 2901–2914, May 2022. doi: 10.1109/JLT.2022.3146157.

[128] N. Andriolli, P. Velha, M. Chiesa, A. Trifiletti, and G. Contestabile, “A directly modulated multiwavelength transmitter monolithically integrated on InP,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 1, pp. 1–6, Jan.-Feb. 2018. doi: 10.1109/JSTQE.2017.2746002.

[129] M. Smit, K. Williams, and J. Van Der Tol, “Past, present, and future of InP-based photonic integration,” *Apl Photonics*, vol. 4, no. 5, May 2019. doi: 10.1063/1.5087862.

[130] R. Stabile, “Towards large-scale fast reprogrammable SOA-based photonic integrated switch circuits,” *Applied Sciences*, vol. 7, no. 9, p. 920, Sep. 2017. doi: 10.3390/app7090920.

[131] W. Yao, "Towards a high-capacity multi-channel transmitter in generic photonic integration technology," Phd Thesis 1 (Research TU/e / Graduation TU/e), Electrical Engineering. ISBN 978-90-386-4194-2 2017.

[132] S. Verghese *et al.*, "Arrays of InP-based avalanche photodiodes for photon counting," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 13, no. 4, pp. 870–886, Jul.-Aug. 2007. doi: 10.1109/JSTQE.2007.904464.

[133] W. Tian, B. Bas, D. Harmsen, K. Williams, and X. Leijtens, "Temperature sensing diode in InP-based photonic integration technology," *IEEE Photonics Journal*, vol. 16, no. 2, pp. 1–8, Apr. 2024. doi: 10.1109/JPHOT.2024.3374266.

[134] O. Çirkinoglu, "Indium-phosphide-based photonic integrated circuits for quantum key distribution," Phd Thesis 1 (Research TU/e / Graduation TU/e), Electrical Engineering. ISBN 978-90-386-6005-9 2024.

[135] W. Guo *et al.*, "Two-dimensional optical beam steering with InP-based photonic integrated circuits," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 19, no. 4, pp. 6 100 212–6 100 212, Jul.-Aug. 2013. doi: 10.1109/JSTQE.2013.2238218.

[136] J. Midkiff, K. M. Yoo, J.-D. Shin, H. Dalir, M. Teimourpour, and R. T. Chen, "Optical phased array beam steering in the mid-infrared on an InP-based platform," *Optica*, vol. 7, no. 11, pp. 1544–1547, Nov. 2020. doi: 10.1364/OPTICA.400441.

[137] A. Pérez-Serrano *et al.*, "Evaluation of radiation hardness of InP-based photonic integrated circuits for space applications," in *International Conference on Space Optics — ICSO 2022*, K. Minoglou, N. Karafolas, and B. Cugny, Eds., vol. 12777, International Society for Optics and Photonics. SPIE, Jul. 2023, p. 127776F. doi: 10.1117/12.2691127.

[138] A. Hänsel and M. J. R. Heck, "Widely tunable laser on an InP photonic integrated circuit," *IEEE Photonics Technology Letters*, vol. 32, no. 2, pp. 105–108, Jan. 2020. doi: 10.1109/LPT.2019.2958711.

[139] T. Merelle. What is a process design kit (PDK)? [Accessed: 2024, July]. [Online]. Available: https://smartphotonics.nl/blog/what-is-a-process-design-kit/

[140] Multi project wafer services. [Accessed: 2024, July]. [Online]. Available: https://www.lionix-international.com/photonics/mpw-services/

[141] A. Zozulia, K. Williams, and Y. Jiao, "Directly modulated lasers on inp membrane platform: design and simulation," in *Proceedings of the Annual Symposium of the IEEE Photonics Benelux Chapter*, 2021, 25th Annual Symposium of the IEEE Photonics Benelux Chapter ; Conference date: 25-11-2021 Through 26-11-2021. [Online]. Available: https://www.aanmelder.nl/ieee-ps-benelux-2021

[142] R. R. Kumar, A. Hänsel, P. Castera, N. Volet, and M. J. R. Heck, "Low-kappa DBR grating filters on an InP generic photonic integration foundry platform," *J. Opt. Soc. Am. B*, vol. 41, no. 4, pp. 1054–1059, Apr. 2024. doi: 10.1364/JOSAB.518800.

[143] C. Chen *et al.*, “Reconfigurable optical interleaver modules with tunable wavelength transfer matrix function using polymer photonics lightwave circuits,” *Optics Express*, vol. 22, no. 17, pp. 19 895–19 911, Aug. 2014. doi: 10.1364/OE.22.019895.

[144] M. Pantouvaki, C. C. Renaud, P. Cannard, M. J. Robertson, R. Gwilliam, and A. J. Seeds, “Fast tuneable InGaAsP DBR laser using quantum-confined stark-effect-induced refractive index change,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 13, no. 5, pp. 1112–1121, Sep.-Oct. 2007. doi: 10.1109/JSTQE.2007.906046.

[145] A. Chraplyvy, “The coming capacity crunch,” in *in Proceedings of the 35th European Conference on Optical Communication ECOC 2009*, Vienna, Austria, 2009, p. Mo1.0.2.

[146] I. P. Kaminow, T. Li, and A. E. Willner, “Digital signal processing (DSP) and its application in optical communication systems,” in *Optical Fiber Telecommunications VIB*, 6th ed. Elsevier Inc, 2013, ch. 5, pp. 163–219. ISBN 9780123969606

[147] A. E. Willner, S. Khaleghi, M. R. Chitgarha, and O. F. Yilmaz, “All-optical signal processing,” *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 660–680, Feb. 2014. doi: 10.1109/JLT.2013.2287219.

[148] W. Liu *et al.*, “A fully reconfigurable photonic integrated signal processor,” *Nature Photonics*, vol. 10, pp. 190–195, Feb. 2016. doi: 10.1038/nphoton.2015.281.

[149] T. Miki, “The potential of photonic networks,” *IEEE Communications Magazine*, vol. 32, no. 12, pp. 23–27, Dec. 1994. doi: 10.1109/35.335996.

[150] Y. Liu *et al.*, “Ultra-low-loss silicon nitride optical beamforming network for wideband wireless applications,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 4, pp. 1–10, Jul.-Aug. 2018. doi: 10.1109/JSTQE.2018.2827786.

[151] D. Pérez *et al.*, “Multipurpose silicon photonics signal processor core,” *Nature communications*, vol. 8, no. 1, p. 636, Sep. 2017. doi: 10.1038/s41467-017-00714-1.

[152] C. Huang *et al.*, “Demonstration of photonic neural network for fiber nonlinearity compensation in long-haul transmission systems,” in *Proc. Optical Fiber Communication Conference, OFC 2020*. Optica Publishing Group (formerly OSA), 2020. doi: 10.1364/OFC.2020.Th4C.6. Paper Th4C.6.

[153] F. Qamar, M. K. Islam, S. Z. Ali Shah, R. Farhan, and M. Ali, “Secure duobinary signal transmission in optical communication networks for high performance & reliability,” *IEEE Access*, vol. 5, pp. 17 795–17 802, Sep. 2017. doi: 10.1109/ACCESS.2017.2752002.

[154] R. Kaur and S. Dewra, “Duobinary modulation format for optical system—a review,” *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 3, no. 8, pp. 11 039–11 046, Aug. 2014. doi: 10.15662/ijareeie.2014.0308012.

[155] A. Lender, “The duobinary technique for high-speed data transmission,” *IEEE Transactions on Communication and Electronics*, vol. 82, no. 2, pp. 214–218, May 1963. doi: 10.1109/TCE.1963.6373379.

[156] J. Sinsky, A. Adamiecki, and M. Duelk, “10-Gb/s electrical backplane transmission using duobinary signaling,” in *2004 IEEE MTT-S International Microwave Symposium Digest (IEEE Cat. No.04CH37535)*, vol. 1, Jun. 2004, pp. 109–112. doi: 10.1109/MWSYM.2004.1335814.

[157] J. Sinsky, M. Duelk, and A. Adamiecki, “High-speed electrical backplane transmission using duobinary signaling,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 1, pp. 152–160, Jan. 2005. doi: 10.1109/TMTT.2004.839326.

[158] J. H. Sinsky, A. Konczykowska, A. L. Adamiecki, F. Jorge, and M. Duelk, “39.4-Gb/s duobinary transmission over 24.4 m of coaxial cable using a custom indium phosphide duobinary-to-binary converter integrated circuit,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, no. 12, pp. 3162–3169, Dec. 2008. doi: 10.1109/TMTT.2008.2007430.

[159] T. LeCroy. The fundamentals of PAM4. [Accessed: 2024, June]. [Online]. Available: https://blog.teledynelecroy.com/2015/07/the-fundamentals-of-pam4.html

[160] W. Han, Y. Wang, and J. Wang, “A 56-Gbps PAM4 amplitude-rectification-based receiver with threshold adaptation and 1-tap DFE,” *IEICE Electronics Express*, vol. 18, no. 17, pp. 1–6, 2021. doi: 10.1587/elex.18.20210302.

[161] V. Houtsma, D. van Veen, and H. Chow, “Demonstration of symmetrical 25 Gb/s TDM-PON with multilevel interleaving of users,” *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 2005–2010, Apr. 2016. doi: 10.1109/JLT.2015.2506140.

[162] M. V. Stephanie, L. Pham, A. Schindler, T. Grasser, M. Waltl, and B. Schrenk, “Photonic neuron with on frequency-domain ReLU activation function,” *Journal of Lightwave Technology*, pp. 1–10, 2024. doi: 10.1109/JLT.2024.3413976.

[163] M. Smit, “New focusing and dispersive planar component based on an optical phased array,” *Electronics Letters*, vol. 24, no. 7, pp. 385–386, 1988. doi: 10.1049/EL:19880260.

[164] T. Segawa *et al.*, “All-optical wavelength-routing switch with monolithically integrated filter-free tunable wavelength converters and an AWG,” *Opt. Express*, vol. 18, no. 5, pp. 4340–4345, Mar. 2010. doi: 10.1364/OE.18.004340.

[165] K. Okamoto, M. Okuno, A. Himeno, and Y. Ohmori, “16-channel optical add/drop multiplexer consisting of arrayed-waveguide gratings and double-gate switches,” *Electronics Letters*, vol. 32, pp. 1471–1472, Aug. 1996. doi: 10.1049/el:19960981.

[166] X. Leijtens, B. Kuhlow, and M. Smit, *Arrayed waveguide gratings*. Germany: Springer, 2006, pp. 125–187. ISBN 978-3-540-31769-2

[167] X. Xiao *et al.*, “Silicon photonic Flex-LIONS for bandwidth-reconfigurable optical interconnects,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 2, pp. 1–10, Mar.-Apr. 2020. doi: 10.1109/JSTQE.2019.2950770.

[168] K. Suzuki *et al.*, “Low-insertion-loss and power-efficient 32×32 silicon photonics switch with extremely high-δ silica PLC connector,” *Journal of Lightwave Technology*, vol. 37, no. 1, pp. 116–122, Jan. 2019. doi: 10.1109/JLT.2018.2867575.

[169] X. Li, W. Gao, L. Lu, J. Chen, and L. Zhou, “Ultra-low-loss multi-layer 8×8 microring optical switch,” *Photonics Research*, vol. 11, no. 5, pp. 712–723, Jan. 2023. doi: 10.1364/PRJ.479499.

List of Publications

Journal Articles

[1] M. V. Stephanie, F. Honz, N. Vokic, W. Boxleitner, M. Waltl, T. Grasser, and B. Schrenk, "SOA-REAM Assisted Synaptic Receptor for Weighted-Sum Detection of Multiple Inputs (Top-Scored Paper)", *Journal of Lightwave Technology*, vol. 41, no. 4, pp.1258–1264, 2023, doi: 10.1109/JLT.2022.3212111.

[2] M. V. Stephanie, L. Pham, A. Schindler, T. Grasser, M. Waltl, and B. Schrenk, "Photonic Neuron with on Frequency-Domain ReLU Activation Function", *Journal of Lightwave Technology*, (accepted).

[3] B. Schrenk and M. V. Stephanie, "ΦPU—A Photonic Processing Unit for Heterogeneous Optical Networks", *Journal of Lightwave Technology*, (accepted).

Conference Contributions

[4] M. V. Stephanie, M. Waltl, T. Grasser and B. Schrenk, “WDM-Conscious Synaptic Receptor Assisted by SOA+EAM”, in *Proc. Optical Fiber Communication (OFC) Conference*. 2022.03.06–2022.03.10, San Diego, United States, pp. 1–3, 2022, paper M1G.2.7.

[5] M. V. Stephanie, F. Honz, N. Vokic, W. Boxleitner, M. Waltl, T. Grasser, and B. Schrenk, “Reception of Frequency-Coded Synapses through Fabry-Perot SOA - REAM Integrating Weighting and Detection Functions”, in *Proc. European Conference on Optical Communication (ECOC) Conference*. 2022.09.18–2022.09.22, Basel, Switzerland, pp. 1–4, 2022, paper Tu5.20.

[6] M. V. Stephanie, L. Pham, A. Schindler, M. Waltl, T. Grasser and B. Schrenk, “All-Optical ReLU as a Photonic Neural Activation Function”, in *Proc. IEEE Photonics Society Summer Topicals Meeting Series (SUM)*. 2023.07.17–2023.07.19, Sicily, Italy, pp. 1–2, 2023, paper MF4.4.

[7] M. V. Stephanie, L. Pham, A. Schindler, M. Waltl, T. Grasser and B. Schrenk, “Optical Neuron with on Frequency-Domain ReLU Activation Function”, in *Proc. Optical Fiber Communication (OFC) Conference*. 2024.03.23–2022.03.28, San Diego, United States, pp. 1–3, 2024, paper M4C.2.

[8] M. V. Stephanie, L. Pham, T. Grasser, M. Waltl, and B. Schrenk, “Optical Neuron with on Frequency - Domain ReLU Activation Function”, in *Proc. European Conference on Optical Communication (ECOC) Conference*. 2024.09.22–2022.03.26, Frankfurt, Germany, pp. 1–4, 2024, paper W2A.44.

Curriculum Vitae

10/2021 – present

Doctoral Candidate
Institute for Microelectronics, TU Wien
Vienna, Austria

10/2018 – 3/2021

Master of Science in Photonics
Thesis titled "Characterisation of Passive Optical Components in Lithium-Niobate-on-Insulator for Photonic Integrated Circuits Applications"
Abbe School of Photonics, Friedrich-Schiller-University of Jena
Jena, Germany

8/2013 – 10/2017

Bachelor of Science in Physics
Thesis titled "Scattering of Electromagnetic Waves on Multiple Spheres"
Faculty of Mathematics and Natural Sciences, Bandung Institute of Technology
Bandung, Indonesia