Ultrafast Photonic Neuromorphic Processing
and Nonlinearity Mitigation in Long-Haul Transmissions by Nonlinear Optical Signal Processing

Yue Tian

A Dissertation
Presented to the Faculty of Princeton University
in Candidacy for the Degree of Doctor of Philosophy

Recommended for Acceptance
by the Department of Electrical Engineering
Adviser: Paul R. Prucnal

June 2014
© Copyright by Yue Tian, 2014.
All rights reserved.
Abstract

As digital electronics and integrated circuits have been dramatically progressed in the past decades, it is increasingly difficult to keep up with Moore’s Law due to the physical limitations of electronics. Recently optical devices are receiving an increased interest as a promising alternative in various areas, including ultrafast signal processing, computing and control systems, because of the low latency and vast bandwidth enabled by nonlinear optical signal processing. The emerging nonlinear material and devices with higher efficiency and smaller footprint have paved the way for on-chip integration and brought nonlinear optical signal processing to practical deployment.

Meanwhile, as an outcome of the marriage of the neuro-ethology drawn from biological neurons with sophisticated modern engineering techniques, neuromorphic processing not only helps in studying biological neural circuits, but also opens up a wide range of applications such as adaptive control, learning, perception, motion control, sensory processing and autonomous robots. Among all the mathematical models drawn from nervous systems, the leaky-integrate-and-fire (LIF) neuron model is the most fundamental and widely used model of biological neurons in theoretical neuroscience for studying complex computation in nervous systems. Its spiking coding and processing mechanism is both computationally efficient and scalable, adopting the best features of both analog and digital computing. Therefore mimicking spike processing with photonics can result in bandwidths that are billions of times higher than biological neurons and substantially faster than electronics. By utilizing a number of nonlinear effects in semiconductor optoelectronic devices and nonlinear fibers, fully functioning photonic neuron prototypes are demonstrated with capability to process optical spikes at the picosecond level. Based on bench-top prototypes, two lightwave neuromorphic circuits are presented as well.

Furthermore, one of the most powerful capabilities of neurons is their ability to learn, which takes place by strengthening synaptic connections in response to spiking
activity. To mimic this process, an optical spiking time dependent plasticity (STDP) device is invented by using nonlinear properties in semiconductor components. With the optical STDP device, for the first time the supervised learning of a photonic neuron is demonstrated, potentially laying the foundation for learning at speeds that are a billion times faster than those of biological neurons.

In addition, the application of nonlinear optical signal processing in telecommunication is studied as well. The huge transmission capacity and ultra-long distance in coherent long-haul transmission systems requires ultrafast signal processing to overcome the linear and nonlinear impairments that occur during transmission. However it is extremely hard for electronics to accomplish real-time processing at such high speeds. With the help of nonlinear optical signal processing, a phase-sensitive boosting (PSB) scheme is proposed using all-optical phase conjugation to mitigate nonlinear impairments and greatly extend the system reach in coherent long-haul optical transmission systems.
Acknowledgements

I would first like to express my gratitude to my research advisor, Prof. Paul R. Prucnal, for sharing his extensive knowledge and insight and for his constant support throughout my graduate studies. Without his direction and encouragement, much of my work would not be possible.

Secondly, I would like to thank all the members of the Lightwave Communications Laboratory for their support and company both inside and outside of the lab. I would like to thank Dr. Mable P. Fok for being a great mentor and for her enormous help. I would also like to thank Dr. Zhenxing Wang and Dr. Yanhua Deng for helping me in the lab. I would also like to thank Dr. Konstantin Kravtsov for helpful discussions during his visits and John Chang for proofreading my thesis. Further I would like to thank Dr. Bhavin J. Shastri, John Chang, Matthew P. Chang, Ben Wu, Alexander N. Tait and Mitchell A. Nahmias for being such great lab mates and friends. With them, I have had countless joyful memories in this amazing group.

Thirdly, I would like to thank all the colleagues at NEC Laboratories America for their support during my internship. I would like to thank Dr. Ting Wang for recruiting me as a summer intern and mentoring me. I would also like to thank Dr. Yue-Kai Huang and Dr. Shaoliang Zhang for their help in the lab, thank Dr. Ezra Ip, Dr. Fatih Yaman and Dr. Akihiro Tanaka for their helpful discussions, thank Dr. Neda Cvijetic for proofreading my thesis.

Fourthly, I especially would like to thank Dr. David Rosenbluth of the Lockheed Martin Advanced Technologies Laboratory for his support and numerous help on the photonic neuron project. I would like to thank Prof. Claire F. Gmachl and Prof. Gerard Wysocki for being my general committee member and providing helpful discussions and directions.

Finally, I would like to thank my parents, Yingzhi Tian and Shihui Wu, for their constant supports and encouragement, without which I could not have the opportu-
nity to pursue my dream. I would also like to thank my wife, Yini Wang, for her faith in me and staying with me across the pacific ocean for nearly five years.
To my wife and parents.
## Contents

Abstract ................................................................. iii  
Acknowledgements .................................................. v  
List of Tables ......................................................... x  
List of Figures ......................................................... xi  

1 Introduction ......................................................... 1  

2 Photonic Neuromorphic Signal Processing: Background and Model 5  
2.1 Introduction ..................................................... 6  
2.2 Neuromorphic processing in electronics and photonics .......... 8  
2.3 Photonic spike processing ...................................... 9  

2.3.1 Spiking signals ............................................... 11  
2.3.2 Spiking processor: computational primitive ............... 12  
2.4 Spiking neuron model .......................................... 16  

3 Photonic Neuron Prototype and Neuromorphic Circuits 19  
3.1 The first photonic neuron (previous work) ...................... 20  

3.1.1 Dendrites—weighting, delay and summation ................. 22  
3.1.2 Soma—temporal integration .................................. 23  
3.1.3 Axon hillock—thresholding and output .................... 27  
3.2 Elemental improvements for enhanced photonic neuron ........ 30  

3.2.1 Versatile and non-inverted integrator ..................... 31
3.2.2 Enhanced thresholder ........................................... 44
3.2.3 Asynchronous spiking photonic neuron ...................... 52

3.3 Lightwave Neuromorphic Circuits ................................. 57
3.3.1 Signal feature recognition based on crayfish tail-flip escape re-
     sponse .............................................................. 58
3.3.2 Timing jitter insensitive photonic logic gate ................. 64
3.4 Conclusion .......................................................... 69

4 Supervised Learning in Photonic Neurons .......................... 72
4.1 Hebbian learning rules and spike timing dependent plasticity .... 73
4.2 Optical spike timing dependent plasticity .......................... 75
4.3 Supervised learning in a photonic neuron based on optical STDP .... 80
4.4 Conclusion .......................................................... 88

5 Phase-Sensitive Boosting for Long-Haul WDM Systems using Opti-
    cal Phase Conjugation ............................................. 89
5.1 Introduction .......................................................... 90
5.2 Principle of digital phase-sensitive boosting technology .......... 92
5.3 Numerical simulation ............................................... 94
5.4 Experimental demonstration ....................................... 97
5.5 Simulation comparison between PSA system and EDFA system with
     and without PSB .................................................. 101
      5.5.1 Analytical Modeling ...................................... 102
      5.5.2 Simulation results and discussions ....................... 108
5.6 Conclusion .......................................................... 111

6 Conclusions .......................................................... 112

Bibliography ........................................................... 116
List of Tables

2.1 Comparison of digital, analog, and hybrid systems . . . . . . . . . . . 12
2.2 Spiking neural automata properties . . . . . . . . . . . . . . . . . . 14
2.3 Photonic elements required by LIF neurons . . . . . . . . . . . . . 15

4.1 Tunability of STDP window . . . . . . . . . . . . . . . . . . . . . . 80

5.1 Fiber parameters used in simulation . . . . . . . . . . . . . . . . . . 104
5.2 Parameters for highly nonlinear fiber used in the simulation . . . . . 106
5.3 Performance of the designed PSA: theory and simulation . . . . . . 108
List of Figures

2.1  Encoding scheme in spiking neural networks .......................... 11
2.2  Schematic and operation of a leaky integrate-and-fire neuron ....... 17

3.1  Block diagram of the first photonic neuron ............................ 21
3.2  The measured SOA response to excitation by multiple pulses ....... 25
3.3  Nonlinear optical loop mirror-based optical thresholder ............. 28
3.4  Transfer function of NOLM-based optical thresholder ............. 29
3.5  Change in an SOA’s gain curve ...................................... 33
3.6  Fiber-to-fiber gain/loss in an SOA as a function of driving current .. 34
3.7  Gain dynamics vs. driving current in SOA ............................ 35
3.8  On-off extinction ratio as a function of SOA driving current ....... 36
3.9  Experimental results of simultaneous operation of gain pumping and depletion ............................................................. 38
3.10 Measured absorption recovery time as a function of EAM bias voltage and optical pump power .......................................................... 40
3.11 Integration response of EAM .............................................. 42
3.12 Time sensitive signal integration in EAM ............................... 43
3.13 Schematic of enhanced NOLM with BPF .............................. 45
3.14 Power transfer function of NOLM ....................................... 47
3.15 Power transfer function of enhanced NOLM with BPF ............... 48
3.16 Power transfer function (low level region) of enhanced NOLM with BPF 49
5.5 Simulation configurations for DC-PSA link and DU-DFA link . . . . . . 103
5.6 Simulation models for non-degenerate PSA with polarization diversity and quantum noise from fiber attenuation . . . . . . . . . . . . . . . 106
5.7 Received signal Q-factor vs. channel launch power for different transmission schemes in single channel operation . . . . . . . . . . . . . . . 109
5.8 Received signal Q-factor vs. channel launch power for different transmission schemes in WDM operation . . . . . . . . . . . . . . . . . . . . . . . . . . 110
Chapter 1

Introduction

After the dramatically rapid development of integrated circuits in the past decades, electronics is approaching its fundamental physical limits. To keep up with Moore’s Law, there is growing interest in optical devices as an alternative in the areas of ultra-fast signal processing and computing. Light signals and optical components have some advantages over electronic counterparts, such as vast bandwidth resource, extremely high capacity and immunity to electromagnetic interference, which make them intrinsically suitable for future signal processing and computing. All-optical processing of high-speed light signals, enabled by nonlinear optical devices, has been receiving increasing recent interests. Progress in nonlinear materials, such as highly nonlinear fibers, photonic-crystal structures, silicon and chalcogenide waveguides, quasi-phase-matched crystals etc., has led to the reduction of the power required for nonlinear optical signal processing down to levels compatible with compact semiconductor lasers, and has paved the way to on-chip integration, which is bringing nonlinear optical signal processing to practical deployment.

At the same time, in an effort to break the limitations inherent in traditional von Neumann architectures, some recent projects in computing have sought more effective signal processing techniques by leveraging the underlying physics of devices
Cognitive computing platforms inspired by biological neural networks could solve unconventional computing problems and outperform current technology in both power efficiency and complexity [8][9]. These novel systems rely on alternative sets of computational principles, including hybrid analog-digital signal representations, co-location of memory and processing, unsupervised learning, and distributed representations of information.

In this thesis, we focus on the optical implementation of the leaky-integrate-and-fire (LIF) model, which is a mathematical model of the spiking dynamics that pervade animal nervous systems, well-established as the most widely used model of biological neurons in theoretical neuroscience for studying complex computation in nervous systems [10]. The concept, coding mechanism and mathematical model of a spiking neuron is introduced in Chapter 2 to explain how a spiking neuron combines the best of both analog and digital worlds, what elements a LIF model needs to function as a computational primitive, as well as current development of neuromorphic processing in electronics and photonics.

Since the first photonic neuron bench-top prototype was presented and demonstrated in our lab years ago [11], many efforts have been made to improve its elemental designs and functionality. New techniques and components have been utilized for the neuron’s key elements, e.g. the temporal integrator and thresholder. Significant improvements for the photonic neuron are shown in Chapter 3 in terms of performance and functionality, as well as two biological-neural-behavior inspired neuromorphic circuits based on our enhanced photonic neurons. A signal feature recognition device inspired by the escape response of a crayfish is demonstrated. By two electro-absorption-modulator- (EAM-) based integrators and one nonlinear-optical-loop-mirror (NOLM) based thresholder, the device is able to be configured to accomplish fast and accurate feature recognition of three-bit input patterns at the picosecond level [12]. Also, based on the gain pumping and gain depletion in
a semiconductor-optical-amplifier (SOA) integrator and a NOLM thresholder, we demonstrated a timing jitter insensitive optical logic gate, which can be reconfigured between AND and NOT gating [13].

Furthermore, one of the most amazing and powerful capabilities of biological neurons is their ability to learn. Learning based on experience is essential for a system to dynamically adapt to an unpredictable or changing environment, or one that is too complex to be characterized a priori. The learning mechanism, called “synaptic plasticity” configures a set of system parameters, such as synaptic weights, based upon the activity of the system in response to its inputs, and as a consequence, the system function changes. Modifying synaptic strength based on correlation with firing is called “Spike Timing Dependent Plasticity” (STDP) [14]. In Chapter 4 we report for the first time the demonstration of supervised learning based on STDP using photonic technology [15], potentially laying the foundation for learning at speeds a billion times faster than biological neurons.

Another area requiring ultrafast signal processing technology is telecommunications. Nowadays coherent detection and digital signal processing (DSP) have been the enabling technologies for improving spectral efficiency in long-haul wavelength-division multiplexing (WDM) transmission. With these sophisticated technologies, the single channel transmission capacity has been pushed to more than 1 Tb/s in long-haul transmissions [16]. However it has been extremely difficult for electronic devices to do real-time processing at such a high operational speed. In Chapter 5 we propose and demonstrate a scheme using all-optical phase conjugation to mitigate the nonlinear impairments in high-speed coherent ultra-long-haul transmission systems [17][18]. The scheme utilizes a phase-conjugated idler signal, which is generated by a nonlinear effect—four-wave mixing (FWM)— and transmitted together with the original signal. After transmission, the signal and idler are jointly detected and processed to suppress nonlinear phase distortion due to long-haul transmission and
increase optical signal-to-noise ratio (OSNR). Both numerical simulation and experimental demonstration show a significant improvement of signal quality and increase of system reach.
Chapter 2

Photonic Neuromorphic Signal Processing: Background and Model

There has been tremendous recent growth in interest in spiking neural networks (SNNs), which code information as spikes or events in time. Spike encoding is widely accepted as the information representation used in the brain, but it has also inspired a new generation of neuromorphic hardware. Although electronics can match biological time scales and exceed them, they eventually reach a bandwidth fan-in trade-off. An alternative hardware instantiation is photonics, which can process highly dynamic information at speeds that electronics could never reach. Correspondingly, processing techniques inspired by biology could compensate for many of the shortcomings that bar digital photonic computing from feasibility, including high defect rates and signal control problems. In this chapter, we summarize previous works on neuromorphic signal processing, properties of photonic spike processing and the mathematical model we used for our photonic neuron.
2.1 Introduction

The brain, unlike the von Neumann architecture found in conventional computers, is very power efficient, extremely effective at certain computing tasks, and highly adaptable to novel situations and environments. While these favorable properties are often largely credited to the unique connectionism of biological nervous systems, it is clear that the cellular dynamics of individual neurons also play an indispensable role in attaining these performance advantages. On the cellular level, neurons operate on information encoded as events or spikes, a type of signal with both analog and digital properties. Spike processing exploits the efficiency of analog signals while overcoming the problem of noise accumulation inherent in analog computation, which can be seen as a first step to attaining the astounding capabilities of bio-inspired processing [19].

Other physical and signal processing features at the single neuron level, including hybrid analog-digital signals, representational interleaving, co-location of memory and processing, unsupervised statistical learning, and distributed representations of information, have been implicated in various positive and conventionally difficult signal processing capabilities. Engineers have tried for nearly half a century to take inspiration from neuroscience by separating key properties that yield performance from simple idiosyncrasies of biology.

If key biological properties are abstracted while others deemed irrelevant, then a bio-inspired engineered system could also incorporate nonbiological properties that may lead to computational domains that are potentially unexplored and/or practically useful. Many microelectronic platforms have attempted to emulate some advantages of neuron-like architectures while incorporating techniques in technology; however, the majority target rather than exceed biological time scales. What kind of signal processing would be possible with a bio-inspired visual front-end that operates ten million times faster than its biological counterpart? Unfortunately, microelectronic
neural networks that are both fast and highly interconnected are subject to a fundamental bandwidth connection-density tradeoff, limiting their speed.

Photonic platforms offer an alternative approach to microelectronics. The high speeds, high bandwidth, and low cross-talk achievable in photonics are very well suited for an ultrafast spike-based information scheme with high interconnection densities. In addition, the high wall-plug efficiencies of photonic devices may allow such implementations to match or eclipse equivalent electronic systems in low energy usage. Because of these advantages, photonic spike processors could access a picosecond, low-power computational domain that is inaccessible by other technologies. Our work aims to synergistically integrate the underlying physics of photonics with bio-inspired spike-based processing. This novel processing domain of ultrafast computing could have numerous applications where temporally precise and robust systems are necessary, including: adaptive control, learning, sensory processing, and autonomous robotics.

Our work on building a photonic neuron will start from a hybrid computational primitive: the spiking neuron, which integrates a small set of basic operations (delay, weighting, spatial summation, temporal integration, and thresholding) into a single device which is capable of performing a variety of computations depending on how its parameters are configured (e.g., delays, weights, integration time constant, threshold). The leaky-integrate-and-fire (LIF) neuron model is a mathematical model of the spiking dynamics which pervades animal nervous systems, and is well-established as the most widely used model of biological neurons in theoretical neuroscience for studying complex computation in nervous systems [10]. LIF neurons have recently attracted the attention of engineers and computer scientists for the following reasons:

- Algorithmic expressiveness: They provide a powerful and efficient computational primitive from the standpoint of computational complexity and computability theory.
• Hardware efficiency/robustness: They are both robust and efficient processing elements. Complex algorithms can be implemented with less hardware.

• Utility in signal processing: There are known pulse processing algorithms for complex signal processing tasks (drawn from neuro-ethology). Pulse processing is already being used for implementing robust sensory processing algorithms in analog very-large-scale integration (VLSI).

2.2 Neuromorphic processing in electronics and photonics

Spike processing algorithms are well understood in a number of important biological sensory processing systems that have found growing use in electronic signal processing applications. The marriage of these physiological principles with engineering [20] [21] [22] not only helps in studying biological neural circuits, but also opens up a wide range of applications such as adaptive control, learning, perception, motion control, sensory processing (vision systems, auditory processors, and olfactory systems), and autonomous robots. For example, using analog VLSI technology, small low-power front-end sensor devices have been implemented that closely replicate the capabilities of the retina and the cochlea [23].

Currently, developing cortically-inspired microelectronic architectures including IBM’s neurosynaptic core [1] [2] and HP’s proposed memristive nanodevices [3] [4] use a dense mesh of wires called a crossbar array to achieve heightened network configurability and fan-in, which are less critical in conventional architectures. These architectures aim to target clock rates comparable to biological time scales rather than to exceed them. At high-bandwidths, however, densely packed electronic wires cease to be effective for communication. Power use increases drastically, and signals quickly attenuate, disperse, or couple together unfavorably, especially on a crossbar
array, which has a large area of closely packed signal wires. In contrast, photonic channels can support the high bandwidth components of spikes without an analogous speed, power, fan-in, and cross-talk trade-off.

Optical neuromorphic technology has been investigated since the early 80’s. Neuron-like devices have been demonstrated using various techniques including holograms/liquid crystals with spatial light modulators [24][25][26], injection locking of lasers [27][28], self-pulsation of semiconductor devices [29], and electron trapping materials [30]. However, only incomplete neurons operating at low speed have been demonstrated with the above approaches and they all rely on electronics to perform key neuronal functions.

2.3 Photonic spike processing

Optical communication has extensively utilized the high bandwidth of photonics, but, in general, approaches to optical computing have been hindered by scalability problems. We hypothesize that the primary barrier to exploiting the high bandwidth of photonic devices for computing lies in the model of computation being used, not solely in the performance, integration, or fabrication of the devices. Many years of research have been devoted to photonic implementation of traditional models of computation, yet neither analog nor digital approaches has proven scalable due primarily to challenges of cascadability and fabrication reliability. Analog photonic processing has found widespread application in high bandwidth filtering of microwave signals, but the accumulation of phase noise, in addition to amplitude noise, makes cascaded operations particularly difficult. Digital logic gates that suppress noise accumulation have also been realized in photonics, but photonic devices have not yet met the extremely high fabrication yield required for complex digital operations, a tolerance requirement that is greatly relaxed in systems capable of biomorphic adaptation. In
addition, schemes that take advantage of the multiple available wavelengths require ubiquitous wavelength conversion, which can be costly, noisy, inefficient, and complicated.

The optical channel is highly expressive and correspondingly very sensitive to phase and frequency noise. Any proposal for a computational primitive must address the issue of practical cascadability, especially if the use of multiple wavelength channels is intended. Our proposed unconventional computing primitive addresses the traditional problem of noise accumulation by interleaving physical representations of information. Representational interleaving, in which a signal is repeatedly transformed between coding schemes (digital-analog) or physical variables (electronic-optical), can grant many advantages to computation and noise properties. For example, a logarithm transform can reduce a multiplication operation to a simpler addition operation. As discussed in the following section 2.3.1, the spiking model found in biology naturally interleaves robust, discrete representations for communication signals with precise, continuous representations for computation variables in order to leverage the benefits of both types of coding. It is natural to deepen this distinction to include physical representational aspects, with the important result that optical noise does not accumulate. When a pulse is generated, it is transmitted and routed through a linear optical network with the help of its wavelength identifier. It is received only once into a fermionic degree of freedom, such as a deflection of carrier concentration in a photonic semiconductor or a current pulse in a photodiode. The soma tasks occur in a domain that is in some ways computationally richer (analog) and in other ways more physically robust (incoherent) to the type of phase and frequency noise that can doom optical computing architectures.

Another major hurdle faced by substantial photonic systems is relatively poor device yield. While it is difficult for digital circuits to build in redundancy overhead, neuromorphic systems are naturally reconfigurable and adaptive. Biological nervous
systems adapt to shunt signals around faulty neurons without the need for external control. The structural fluidity of neural algorithms will itself provide an effective resistance against fabrication defects.

2.3.1 Spiking signals

Figure 2.1: Spiking neural networks encode information as events in time rather than bits. The time at which a spike occurs is analog while its amplitude is digital, the signals use a mixed-signal or hybrid encoding scheme.

On the cellular level, the brain encodes information as events or spikes in time [31], “hybrid signals” with both analog and digital properties as illustrated in Figure 2.1. Spike processing has evolved in biological (nervous systems) and engineered (neuromorphic analog VLSI) systems using LIF neurons as processing primitives that encode data in the analog timing of spikes as inputs. Through its fine grained interleaving of analog and digital processing, it provides the means for achieving scalable high bandwidth photonic processing by overcoming both the analog noise accumulation problem [19] and the bandwidth inefficiency of digital processing. These unique
Properties of the spike processing model of computation enable a hybrid analog and
digital approach that will allow photonics hardware to scale its processing complexity
and efficiency. The foregoing discussion is summarized in Table 2.1, which compares
digital, analog, and hybrid systems in terms of robustness, expressiveness, power effi-
ciency, and the characteristic property that determines usable bandwidth. Spiking
signals, which in addition to being inherently advantageous, carry information in a
natural and accessible fashion that forms the very foundation of some of the astound-
ing capabilities of systems studied in neuroscience.

Table 2.1: Comparison of digital, analog, and hybrid systems in terms of robustness,
expressiveness, power efficiency, and bandwidth limiter

<table>
<thead>
<tr>
<th></th>
<th>Digital</th>
<th>Analog</th>
<th>Spiking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robustness</td>
<td>√</td>
<td>×</td>
<td>√</td>
</tr>
<tr>
<td>Expressiveness</td>
<td>×</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Power Efficiency</td>
<td>×</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Bandwidth Limiter</td>
<td>Clock speed</td>
<td>Minimum SNR</td>
<td>Pulse width</td>
</tr>
</tbody>
</table>

2.3.2 Spiking processor: computational primitive

Like a gate in a digital computer, a spike processor serves as the basic unit of larger
interconnected networks that can perform more complex computations. There are
five key computational properties to consider: (1) integration—the ability to sum
and integrate both positive and negative weighted inputs over time; (2) threshold-
ing—the ability to make a decision whether or not to send a spike; (3) reset—the
ability to have a small refractory period during which no firing can occur immedi-
ately after a spike is released; (4) pulse generation—the ability to generate new pulses;
and (5) adaptability—the ability to modify and regulate response properties on slow
timescales based on statistical properties of environmental inputs and/or training.

These five properties all play important roles in the emulation of the widely ac-
cepted neurocomputational primitive—the LIF neuron. This model is more compu-
tionally powerful than either the rate or the earlier perceptron models [32], and can
serve as the baseline unit for many modern cortical algorithms [33] [34] [35]. Although
not every property is needed to do constrained useful computations, each one serves
an important purpose to assure a high level of richness and robustness in the overall
computational repertoire.

Integration is a temporal generalization of summation in older perceptron-based
models. Spikes with varying amplitudes and delays arrive at the integrator, changing
its state by an amount proportional to their input amplitudes. Excitatory inputs
increase its state, while inhibitory inputs deplete it. The integrator is especially
sensitive to spikes that are closely clustered in time or with high amplitudes, the
basic functionality of temporal summation. Eventually, without any inputs the state
variable will decay to its equilibrium value. Both the amplitude and timing play an
important role for integration.

Thresholding determines whether the output of the integrator is above or below
a predetermined threshold value, \( T \). The neuron makes a decision based on the state
of the integrator, firing if the integration state is above the threshold. Thresholding
is the center of nonlinear decision making in a spiking system, which reduces the
dimensionality of incoming information in a useful fashion and plays a critical role
in cleaning up amplitude noise that would otherwise cause the breakdown of analog
computation.

The reset condition resets the state of the integrator to a low, rest value immedi-
ately after a spike processor fires, causing a refractory period in which it is impossible
or difficult to cause the neuron to fire again. It plays the same role in time as the
thresholder does for amplitude, cleaning up timing jitter and preventing the temporal
spreading of excitatory activity while putting a bandwidth cap on the output of a
given spiking unit.
Pulse generation refers to the ability for a system to spontaneously generate pulses. If pulses are not regenerated as they travel through systems, they will eventually be lost in noise. A system with this property can generate pulses without the need to trigger on input pulses whenever the integrator’s state variable reaches the threshold, $T$.

Adaptability is the network’s ability to change to better suit changing environmental and system conditions. Adaptation of network parameters typically occurs on time scales much slower than spiking dynamics and can be separated into either unsupervised learning, where alterations in overall signal statistics cause automatic adjustments, or supervised learning, where changes are guided based on the behavior of the system compared to a desired behavior presented by a teacher. Because of the extreme reconfigurability and fluidity of massively parallel neural networks, adaptation rules are necessary to stabilize the system’s structure and accomplish a desired task. Adaptation also corrects catastrophic system alterations by, for example, routing signals around a failed neuron to maintain overall process integrity.

<table>
<thead>
<tr>
<th>Integration</th>
<th>Temporally sum weighted inputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thresholding</td>
<td>Fire one spike when integrator state exceeds some level</td>
</tr>
<tr>
<td>Reset</td>
<td>Return integrator state to rest immediately following a spike</td>
</tr>
<tr>
<td>Pulse generation</td>
<td>Introduce a new spike into the network</td>
</tr>
<tr>
<td>Adaptability</td>
<td>Modify behavioral parameters based on input statistics</td>
</tr>
</tbody>
</table>

Table 2.2: Spiking neural automata properties

The characteristics mentioned above are summarized in Table 2.2. In summary, (1) and (4) provide key properties of a system that is the basis of asynchronous communication, while (2) and (3) clean up amplitude and temporal noise, respectively, to allow for cascadability. Any processor designed to closely emulate LIF neurons should have all five properties.
In photonics, the challenge of creating a flexible, scalable, and efficient native hardware implementation of the spike processing model of computation lies in the design of the critical, programmable component devices of the LIF neuron. The ability of the LIF to perform matched filtering, dimensionality reduction, evidence accumulation, decision making, and communication of results requires analog, tunable component hardware elements as described in Table 2.3. These parts must perform adjustable operations and demonstrate configurable interconnectivity between large numbers of LIF neurons.

Table 2.3: Photonic elements required by LIF neurons

<table>
<thead>
<tr>
<th>Challenges</th>
<th>Required analog control</th>
<th>Photonic device candidates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matching filtering</td>
<td>Connectivity weights and delays</td>
<td>Amplifiers, tunable microring resonators</td>
</tr>
<tr>
<td>Dimensionality reduction</td>
<td>Summation of input lines</td>
<td>Couplers, waveguides</td>
</tr>
<tr>
<td>Evidence accumulation</td>
<td>Temporal integration constant</td>
<td>Photodiodes</td>
</tr>
<tr>
<td>Decision making</td>
<td>Adjustable thresholding</td>
<td>Laser bias; saturable absorber devices</td>
</tr>
</tbody>
</table>

The weights and delays of the matched filter are applied at the interconnection of two LIF neurons and form a complicated junction that acts as the primary communication medium between the spiking elements. These interconnection controls are a key feature of the LIF and must be capable of rapid read and write reconfigurability and adjustability either through direct programming or through learning algorithms like spike timing dependent plasticity (STDP).

In the next section, we will review the LIF neuron model which is a mathematical model of the spiking neuron primitive.
2.4 Spiking neuron model

Our photonic neurons are based upon a well-studied and paradigmatic example of a hybrid computational primitive: the spiking neuron. Studies of morphology and physiology have pinpointed the LIF model as an effective spiking model to describe a variety of different biologically observed phenomena [10]. From the standpoint of computability and complexity theory, LIF neurons are powerful and efficient computational primitives that are capable of simulating both Turing machines and traditional sigmoidal neural networks [36]. These units perform a small set of basic operations (delaying, weighting, spatial summation, temporal integration, and thresholding) that are integrated into a single device capable of implementing a variety of processing tasks, including binary classification, adaptive feedback, and temporal logic.

The basic biological structure of a LIF neuron is depicted in Figure 2.2(a). It consists of a dendritic tree that collects and sums inputs from other neurons, a soma that acts as a low pass filter and integrates the signals over time, and an axon that carries an action potential, or spike, when the integrated signal exceeds a threshold. Neurons are connected to each other via synapses, or extracellular gaps, across which chemical signals are transmitted. The axon, dendrite, and synapse all play an important role in the weighting and delaying of spike signals.

According to the standard LIF model, neurons are treated as an equivalent electrical circuit. The membrane potential $V_m(t)$, the voltage difference across their membrane, acts as the primary internal (activation) state variable. Ions that flow across the membrane experience a resistance $R = R_m$ and capacitance $C = C_m$ associated with the membrane. The soma is effectively a first-order low-pass filter, or a leaky integrator, with the integration time constant $\tau_m = R_mC_m$ that determines the exponential decay rate of the impulse response function. The leakage current through $R_m$ drives the membrane voltage $V_m(t)$ to 0, but an active membrane pumping current counteracts it and maintains a resting membrane voltage at a value of $V_m(t) = V_L$. 
Figure 2.2: (a) Schematic and (b) operation of a leaky integrate-and-fire neuron. Weighted and delayed input signals are summed onto the soma, which integrates them together and makes a spike or no-spike decision. The resulting spike is sent to other neurons in the network.

Figure 2.2(b) shows the standard LIF neuron model \([36]\). A neuron has: (1) \(N\) inputs which represent induced currents through input synapses \(\sigma_j(t)\), that are continuous time series consisting either of spikes or continuous analog values; (2) an internal activation state \(V_m(t)\); and (3) a single output state \(O(t)\). Each input is independently weighted by \(\omega_j\) and delayed by \(\tau_j\) resulting in a time series that is spatially summed (summed pointwise). This aggregate input electrical cur-
rent, \( I_{\text{app}}(t) = \sum_{j=1}^{n} \omega_j \sigma_j(t - \tau_j) \). The result is then temporally integrated using an exponentially decaying impulse response function resulting in the activation state \( V_m(t) = V_L e^{\frac{t-t_0}{\tau_m}} + \frac{1}{C_m} \int_{0}^{t-t_0} I_{\text{app}}(t - s) e^{\frac{s}{\tau_m}} \, ds \), where \( t_0 \) is the last time the neuron spiked. If \( V_m(t) \geq V_{\text{thresh}} \), then the neuron outputs a spike, \( O(t) = 1 \), and \( V_m(t) \) is set to \( V_{\text{reset}} \). After issuing a spike, there is a short period of time, the refractory period, during which it is difficult to issue another spike; that is, if \( O(t) = 1 \) then \( O(t + \Delta t) = 0 \), \( \Delta t \leq T_{\text{refract}} \). Consequently, the output of the neuron consists of a continuous time series comprised of spikes.

The parameters determining the behavior of the device are: the weights \( \omega_j \), delays \( \tau_j \), threshold \( V_{\text{thresh}} \), resting potential \( V_L \), refractory period \( T_{\text{refract}} \), and the integration time constant \( \tau_m \). There are three influences on \( V_m(t) \): passive leakage of current, an active pumping current, and external inputs generating time-varying membrane conductance changes. These three influences are the three terms contributing to the differential equation describing \( V_m(t) \) as

\[
\frac{dV_m(t)}{dt} = \frac{V_L}{\tau_m} - \frac{V_m(t)}{\tau_m} - \frac{1}{C_m} I_{\text{app}}(t) \quad (2.1)
\]

if \( V_m(t) > V_{\text{thresh}} \) then release a pulse and set \( V_m(t) \rightarrow V_{\text{reset}} \).
Chapter 3

Photonic Neuron Prototype and Neuromorphic Circuits

In this chapter, the first photonic neuron bench-top prototype is presented and demonstrated by my labmates. It emulates spiking *leaky-integrate-and-fire* (LIF) behavior with ultrafast photonic components. Like its biological counterpart, the photonic neuron consists of several elements, i.e. connectivity weights and delays, summation, temporal integration and thresholding. All these elements are realized in ultrafast photonic components, which are capable of processing picosecond “spikes”—optical pulses. Then we have made several efforts to improve each element to enhance the photonic neuron’s performance, functionality and versatility. Although this fiber-based photonic neuron is still bulky, power hungry and inefficient, it demonstrates a variety of important hybrid spike processing functions. In addition, based on the enhanced photonic neurons several prototypical lightwave neuromorphic circuits are demonstrated as well.
3.1 The first photonic neuron (previous work)

As was discussed in Section 2.3, spike processing is a hybrid of analog and digital processing in the sense that the output of a neuron is binary in amplitude, while processing in the neuron itself is analog: it does analog weighting and summation/subtraction of the input signals. This type of processing has evolved in biological systems (nervous systems) in order to overcome the problem of noise accumulation inherent in purely analog computation [19]. Thus all information passed down to the following neuron in the network is contained only in the presence or absence of spikes, but not in their shape or amplitude. Only timing of spikes provides the means of encoding information in spike processing. Spike processing algorithms are well understood in a number of important biological sensory processing systems and are finding growing use in signal processing applications [36]. From the standpoint of computability and complexity theory, integrate-and-fire neurons are powerful computational primitives that are capable of simulating both Turing machines and traditional neural networks [36].

The standard LIF model of a neuron has the following properties:

1. A neuron has \( N \) inputs \( \sigma_i(t) \); it has an internal activation state \( V_m(t) \); and it has a single output state \( O(t) \). At rest, the internal state is actively maintained at \( V_{\text{rest}} \).

2. Inputs \( \sigma_i(t) \) are continuous time series consisting of either spikes or continuous analog values.

3. Inputs are weighted by \( w_i \) and delayed by \( \delta_i \) to give \( w_i \sigma_i(t - \delta_i) \).

4. Delayed and weighted input time series are spatially integrated through point-wise summation, \( \sum_{i=1}^{N} w_i \sigma_i(t - \delta_i) \).
5. The activation state $V_m(t)$ is an exponentially weighted temporal integration of the spatially summed input time series. 

$$V_m(T) = V_{rest} - \frac{1}{C_m} \int_{-\infty}^{T} I(t)e^{-\frac{T-t}{\tau_m}} dt,$$

where $\tau_m$ is the integration time constant, $I(t) = V_m(t) \sum_{i=1}^{N} w_i \sigma_i(t + \delta_i)$ is the electrical current induced by the aggregated input, and $C_m$ is the neuron capacitance.

6. The moment the magnitude of the temporally integrated signal drops below the threshold, then the neuron outputs a spike, $O(t) = 1$ if $|V_m(t)| < |V_{thresh}|$.

7. After issuing a spike there is a short period of time, the refractory period, during which no other spikes can be issued, if $O(t) = 1$ then $O(t - \Delta t) = 0$, $\Delta t \leq T_{refract}$.

8. Output of the neuron consists of a continuous time series of spikes.

The parameters determining the behavior of the device are: $w_i$, $\delta_i$, $V_{thresh}$, $V_{rest}$, $T_{refract}$, and the integration time constant $\tau_m$.

The functional architecture of the first photonic LIF neuron [11] shown in Figure 3.1 consists of three processing blocks: dendritic tree—passive weighting, delay, and summation of inputs; soma—temporal integration; axon hillock—thresholding.

Figure 3.1: Block diagram of the photonic neuron. G: variable attenuator, T: variable delay line, SOA: semiconductor optical amplifier, HD fiber: heavily GeO$_2$-doped fiber.
3.1.1 Dendrites—weighting, delay and summation

The favorable properties of biological neural systems (e.g. brains) are largely credited to the tremendous interconnectivity between neurons. Therefore the “dendrites”, as the input ports of a neuron, have to be high efficient at delivering and combining a huge amount of “spikes”. The high speed, high bandwidth, and low cross-talk achievable in photonics are intrinsically suited for an ultrafast spike-based information scheme with high interconnection densities. Optical fibers are widely used in fiber-optic communications and the best medium to guide light signal in terms of minimizing attenuation. They are lightweight, compact and flexible. Light signals at different wavelengths can be transmitted within a single fiber with little crosstalk, and are immune to outside electromagnetic interference as well. The total supported waveband with ultra-low loss is more than one hundred nanometers around 1550 nm. Compared with their electronic counterparts—electrical wires and cables, optical fibers are inherently suitable for the high interconnection density required by neural systems. Therefore the best candidates for the implementation of the dendritic tree of a photonic neuron are optical fibers and optical fiber-based components.

Besides propagating “spikes” to the soma, dendrites impose different weights and time delays onto these “spikes” from different sources. Before accessing the dendrites the “spikes” have identical amplitude, and in each dendritic branch their amplitude is imposed with specific weight value, which corresponds to how much contribution though this dendritic branch the “spikes” are going to make in the next temporal integration operation. In our bench-top prototype, weighting is simply realized by an optical attenuator, which constantly attenuates the output light power by a certain amount. The weight value—attenuation can be mechanically or electrically tuned. The tuning range can be from 0 to more than 60 dB ($10^{-6}$). Other than the passive attenuators, sometimes we also use electric-optical modulators (EOMs) to generate attenuation to the optical “spikes”, since EOMs have much faster response so as to
act as dynamic weighting units, as in the supervised learning demonstration presented in Section 4.3.

Time delays to the “spikes” are applied by variable optical delay lines, which change the light path length to generate variable time delay. Time delay values on the dendritic branches directly affect the timing relation between the “spikes”, which consequently influences the results of the temporal integration happening in the soma.

Since optical pulses at different wavelengths do not interfere with each other, summation of the optical pulses in multiple fibers (dendritic branches) can be accomplished by directly coupling them into one single fiber. Thus a passive optical coupler is able to realize the summation function.

3.1.2 Soma—temporal integration

As presented at the beginning of this section, the standard LIF neuron model treats biological neurons as electrical devices, which have their membrane potential, i.e. the voltage between the neurons body and the outside, $V_m(t)$, as the primary internal state variable governing the soma’s functionality. The passive electrical properties of the membrane can be modeled as an $RC$ circuit with $R$ referring to the resistance of the membrane and $C$ referring to the capacitance associated with the membrane. Thus the membrane is effectively a first-order low-pass filter, or a leaky integrator with the time constant $\tau_m = R_mC_m$. A leakage current through $R_m$ drives the membrane voltage $V_m(t)$ to 0, but an active membrane pumping current counteracts it and maintains a resting membrane voltage at a value of $V_m(t) = V_{rest}$. Consequently, there are three influences on $V_m(t)$: passive leakage of current; an active pumping current; and external inputs generating time varying membrane conductance changes $\sigma(t)$, which help to “discharge” the neuron. These three influences are the three terms
contributing to the differential equation describing $V_m(t)$ in Equation 3.1 below.

$$\frac{dV_m(t)}{dt} = \frac{V_{rest}}{\tau_m} - \frac{V_m(t)}{\tau_m} + \frac{1}{C_m} V_m(t) \sigma(t)$$  \hspace{1cm} (3.1)

$$\frac{dN'(t)}{dt} = \frac{N_{rest}}{\tau_e} - \frac{N'(t)}{\tau_e} + \frac{\Gamma a}{E_p} N'(t)I(t)$$  \hspace{1cm} (3.2)

Similarly, the gain dynamics of a short semiconductor optical amplifier (SOA) are governed by the Equation 3.2 \[37\]. The SOA’s primary internal state variable is the carrier density above transparency $N'(t) = N(t) - N_0$ where $N(t)$ is the actual carrier density and $N_0$ is the carrier density at transparency. Again, there are three contributors to changing $N'(t)$: a passive leakage due to spontaneous light emission leading to the carrier decay; an active pumping provided by the driving current of the SOA; and stimulated light emission due to the neuron inputs, which also “discharges” the neuron reducing its state variable $N'(t)$. It is remarkable that the electrical model of membrane voltage is essentially identical to the optical model of SOA carrier density. The integration constant of the photonic neuron, $\tau_e$, is equal to the carrier lifetime, and the stimulated emission term in Equation 3.2 depends on the total input light intensity $I(t)$, mode confinement factor $\Gamma$, differential gain coefficient $a$, and photon energy $E_p$. Due to the direct mapping of the electrical model of membrane voltage and the optical model of SOA carrier density, SOA device physics can be used as the analog implementation of a leaky integrator in the optical implementation of the LIF neuron. Since the carrier lifetime in modern SOAs can be as short as 10 ps, photonic neurons using this integration technique may be more than $10^8$ times faster than their biological counterparts.

Temporal integration in the SOA can be experimentally observed by studying the SOA gain dynamics under excitation by a series of input pulses with different inter-arrival times. The effects of the pulses which arrive within the integration window

<table>
<thead>
<tr>
<th>Activation</th>
<th>Active pumping</th>
<th>Leakage</th>
<th>External input</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{dV_m(t)}{dt}$</td>
<td>$V_{rest}$</td>
<td>$- \frac{V_m(t)}{\tau_m}$</td>
<td>$\frac{1}{C_m} V_m(t) \sigma(t)$</td>
</tr>
<tr>
<td>$\frac{dN'(t)}{dt}$</td>
<td>$N_{rest}$</td>
<td>$- \frac{N'(t)}{\tau_e}$</td>
<td>$\frac{\Gamma a}{E_p} N'(t)I(t)$</td>
</tr>
</tbody>
</table>
should add up, while pulses which are far apart should affect the gain independently. Gain measurements were performed by scanning the delay between the input signal, consisting of 16 pulses per each 800 ps, and the sampling pulse stream. The results of the measurement are shown in Figure 3.2.

Figure 3.2: The measured SOA response to excitation by multiple pulses. (a) Sampled SOA gain dynamics; (b) Oscilloscope trace of the control pulse sequence. The SOA recovery time (integration time constant) is 180 ps, resting potential level equals 43 fJ.

The lower plot obtained with a 30-GHz bandwidth sampling oscilloscope represents a sequence of input pulses. Note that on the plot the pulses are stretched in time compared with their original width of 3 ps due to the limited bandwidth of the oscilloscope. The upper plot shows output sampling pulse power vs. relative delay between the signal and the sampling stream. Both plots have the same X-axis range and scale. A direct correspondence between the two is apparent: each input pulse leads to a drop of the SOA gain that gradually recovers until the next pulse comes. If
input pulses are close to each other as in the middle of the plot, there is not enough
time for the gain to recover so the output sampling power remains small, while an
isolated pulse around 700 ps acts after the gain is recovered such that the gain drop is
small. A series of input pulses creates a saw-tooth-like gain dynamics picture shown
in Figure 3.2(a). The two main parameters of integration, the “resting potential”,
i.e. the maximum SOA output if no input signal is present, and the SOA integration
time \(\tau_e\), are measured by studying gain recovery after a single input pulse. They are
equal to 43 fJ and 180 ps respectively. It is also important that the SOA integration
time can be adjusted by changing the SOA pump current. For our SOA (Alcatel
A1901SOA), \(\tau_e\) was changing from 100 ps at 170 mA to 300 ps at 70 mA.

In order to use the magnitude of SOA carrier density as an input to later stages of
processing, it must be converted into a pulse intensity magnitude. Gain sampling is
used as the means of converting SOA carrier density to pulse intensity. A low energy
external pulse train at another wavelength \(\lambda_1\) as shown in Figure 3.1 different from
the wavelengths of the input signal, provides a sampling input to the SOA. Since the
energy of the sampling pulses is small, their effect on the SOA gain is negligible in
comparison to the effect of the input signals. As both the sampling pulses and the
input signal pass through the SOA simultaneously, the amplification of the sampling
pulses provides a measure of SOA carrier density primarily affected by the input
signals. A band-pass spectral filter at the SOA output passes only the sampling
pulses to the thresholder. This process regenerates the optical signal since it is the
sampling pulse train, rather than transformed input signals. It is worth noting that
the output of the integrator, i.e. the intensity magnitude of the amplified sampling
pulse train, is inversely proportional to the integrated result of input optical pulses,
as shown in the subset of Figure 3.1. For example, the larger the integration result
is, the smaller the magnitude output pulse train has, and vice versa.
3.1.3 Axon hillock—thresholding and output

The third important element of the photonic neuron is the axon hillock, which performs thresholding and outputs spikes. When the integrated signal from the temporal integration in soma exceeds a threshold, the axon hillock sends a spike to the output axon; otherwise it remains silent. An effective technology for all-optical thresholding of ultrafast pulses is the well-known nonlinear optical loop mirror (NOLM) [38]. These devices use asymmetrical nonlinear phase shifts in highly nonlinear fiber (HNF) to achieve a polynomial power transfer function (power out vs. power in). A nonlinear phase shift is induced by self-phase modulation (SPM) [39] in the HNF, which causes intensity-dependent interference in a joining coupler. The NOLM is a useful device for optical gating, wavelength conversion and optical code division multiple access networks [38] [40] [41] [42] [43]. Modified versions of the NOLM have been introduced with in-loop amplifiers or asymmetric loss [44] and an in-loop directional attenuator [45]. However, these schemes are quite bulky and have high latency because they typically consist of kilometers of nonlinear fiber. In addition, research works have been conducted on improving the NOLM by changing the parameters of the fiber through the use of special devices or fiber twisting [41], [45], which complicate the settings and destroy the simplicity of the basic NOLM. Among the various modified versions of NOLMs, [40] demonstrates a Ge-doped NOLM with a low requirement on the nonlinear phase shift, allowing the length of the nonlinear element to be shortened (only 15-m of highly Ge-doped nonlinear fiber are used). A tunable directional attenuator is also employed to allow precise and independent control of the amplitudes of the interfering components [40]. Therefore we decide to take this scheme of NOLM as an all-optical thresholder as the axon hillock of our photonic neuron.

In general, the NOLM-based thresholder is a Sagnac interferometer that utilizes SPM to generate an intensity-dependent nonlinear phase shift in order to change the transmittance $T$, which is the ratio of output power to the launched power:
$T = P_{out}/P_{in}$. This SPM effect causes the thresholder to behave nonlinearly with input power. The setup is shown in Figure 3.3. The input pulse is split by a 90/10 coupler into two counter-propagating pulses, one with 90% of the original power, and one with 10% of the original power. The size of the arrows in Figure 3.3 represents the amplitude of the pulse the arrow describes. When each pulse reaches the nonlinear fiber, it incurs a phase shift proportional to its power. Thus, the stronger pulse has a significantly greater phase shift than the weaker pulse. The tunable isolator acts as a directional attenuator, allowing the weak pulse to pass and attenuating the counter-propagating strong pulse. The tunable isolator is tuned to balance the power of the counter-propagating pulses at the coupler output. Thus, when the two pulses meet at the coupler after completing the loop, they have the same power but different phases. The polarization controller allows for fine adjustment of the relative phases of the pulses, according to the Jones matrix. The interference is such that for low powers, they interfere destructively at the output, while for high powers (above the threshold level), they interfere constructively at the output.

Figure 3.3: The setup of a Ge-doped-fiber-based nonlinear optical loop mirror.
To measure the transfer characteristics of the optical thresholder, an optical pulse train at 1550 nm with a full width at half maximum (FWHM) of 2.2 ps is generated by an erbium-doped fiber-based mode-locked laser (MLL). An erbium-doped fiber amplifier (EDFA) with a maximum output power of about 300 mW is directly connected to the input of the thresholder, so as to vary the input optical pulse power into the thresholder by tuning the pump current of the EDFA. The measured and simulated transfer function curves are plotted in Figure 3.4. Figure 3.4 shows a highly nonlinear behavior of the transfer function at low powers (normalized input power < 1), which is exploited for thresholding.

![Figure 3.4: The measured and simulated transfer function of the NOLM-based thresholder.](image)

Due to the utilized mechanism of integration with SOA gain sampling, the result of the thresholding is the absence of a spike when inputs exceed the threshold, and the presence of spikes when input is below threshold. Since this is inverted with respect to the output of the integrate-and-fire model, the inversion stage and another
thresholding for final signal standardization are needed after the optical thresholder. The details of the additional inversion and thresholding stages can be found in [46].

3.2 Elemental improvements for enhanced photonic neuron

Although the first bench-top photonic neuron prototype [11] has realized the functions of essential elements of a neuron, there are some drawbacks stemming from the mechanism of integration with SOA gain sampling. For instance, a biological neuron takes two types of inputs—excitatory and inhibitory, which have opposite impacts on the temporal integration result. However, by using the gain depletion effect in an SOA, the photonic neuron can only take in excitatory input. To support both excitatory and inhibitory input, two adjacent SOA-based integrators have to be used, as in [46]. Even without inhibitory inputs, the inverted presentation of the temporal integration result due to the gain depletion mechanism requires an additional inversion and thresholding stage, which complicates the prototype’s setup. Moreover, the gain sampling method imposes a timing limitation on output optical spikes. Since the output pulses are actually a sampling pulse train whose intensity magnitude was processed by the integrator and thresholder, the timing positions of output pulses are always synchronized with the sampling pulse train instead of input pulses to the neuron. Therefore, the large bandwidth of spiking processing enabled by the spikes’ analog timing position is somewhat compromised by the digital timing of the sampling pulse train. In this section, we propose and demonstrate new techniques [47] [48] [49] [48] to overcome these problems to improve the functionality and reduce the system complexity. In addition, an enhanced optical thresholder scheme [50] is also presented to improve the thresholding performance for photonic neurons.
3.2.1 Versatile and non-inverted integrator

Since optical input pulses deplete the SOA’s carrier density (gain depletion), an SOA-based integrator produces inverted integration results, which is not favorable for photonic neurons. However, we have not fully exploited the carrier dynamic characteristics yet. Besides gain depletion, the gain pumping effect (opposite to the gain depletion effect) utilizes shorter wavelength pump light, i.e. in the S-band (1460–1530 nm), to boost the carrier density in a C-band (1530–1565 nm) SOA, which increases its gain to C-band incident light. It has been demonstrated for various applications, e.g. S- to C/L-band wavelength conversion [51], carrier-recovery time reduction [52] and increasing saturation power [53]. If S-band optical pulses can act as excitatory input and C-band optical pulses as inhibitory input to boost and deplete the SOA’s gain respectively, a single SOA will be able to integrate both types of input spikes and no additional integrator or inversion stage is needed.

Firstly, we analyze the gain dynamics and investigate the properties of gain pumping and depletion in a C-band SOA using both S-band and C-band pump signals. The gain dynamics of an SOA with both gain-depletion and gain-pumping pumps is governed by the rate equation as follows [47][54]:

\[
\frac{dN'(t)}{dt} = \frac{N'_{\text{rest}}}{\tau} - \frac{N'(t)}{\tau} - \frac{\Gamma_d a_d}{E_d} N'(t) I_d(t) + \frac{\Gamma_p a_p}{E_p} N'(t) I_p(t)
\]

(3.3)

where \(N'(t)\) is the carrier density above transparency, \(N'_{\text{rest}}\) is the resting carrier density without any incident pump light, \(\tau\) is the carrier lifetime, \(I(t), \Gamma, a\) and \(E\) represent the input light intensity, the mode confinement factor, the differential gain coefficient, and the photon energy, respectively. The suffix \(d\) or \(p\) stands for gain depletion or pumping.
In an SOA, a wavelength occurs where the material has zero gain to the incident light and consequently the SOA neither amplifies nor consumes the incident optical signal. It is called the material transparency wavelength $\lambda_{tr-m}$, and it is a critical boundary between the gain pumping and depletion. The light at a wavelength longer than $\lambda_{tr-m}$ consumes the SOA’s gain and is amplified (gain depletion), while shorter-wavelength light is absorbed by the SOA and converted into excited carriers. Therefore by carefully picking the wavelength we can choose which gain dynamic effect we would like to use. It is also found that $\lambda_{tr-m}$ is not a fixed number. Instead it varies with carrier density in the gain material. As shown in Figure 3.5, $\lambda_{tr-m}$ is dependent on carrier density, and higher carrier density causes a blue-shift of the gain spectrum to shorter wavelengths. Due to the coupling and waveguide loss, there is a more practical parameter—device transparency wavelength $\lambda_{tr-d}$, where the fiber-to-fiber gain is zero. It is obvious that $\lambda_{tr-d}$ is directly related to and somewhat longer than $\lambda_{tr-m}$. So we experimentally investigate the $\lambda_{tr-d}$ and gain characteristics of a C-band SOA.

For a C-band SOA, an S-band laser, with wavelengths mostly shorter than 1500 nm, is required to generate incident light for gain pumping. Unfortunately, most off-the-shelf lasers at this waveband are pump lasers for EDFA, which do not have a wide tuning range like C-band lasers. Therefore instead of sweeping the laser wavelength across the $\lambda_{tr-d}$, we use a fixed 1479-nm laser, and by tuning the driving condition of the SOA we sweep its transparency wavelength around 1479 nm. In this way, we can learn the gain characteristics of the SOA without a tunable laser. The experiments are performed on a commercial pigtailed SOA (Kamelian NL-L1-C-FA-1550). First, we measure the SOA operating condition when its $\lambda_{tr-d}$ is at 1479 nm. This can be obtained by monitoring the SOA gain at 1479 nm as we tune its driving current. It is worth noting that the input power has to be weak enough, $-24$ dBm in our experiment, so that the carrier density is not severely affected by the incident 1479-
Figure 3.5: Illustration of the change in gain curve of a forward-biased SOA when the carrier density is changed. Dashed line: high carrier density. Solid line: low carrier density.

nm light. The result is shown in Figure 3.6, which indicates there is no gain or loss to the 1479-nm incident light when the driving current is approximately 70 mA. With low driving current (<70 mA), 1479 nm falls in the absorption region, while it sits in the gain region with high driving current (>70 mA). The SOA gain begins to saturate when the driving current exceeds 100 mA. The transparency to the incident light implies that 1479 nm is the device transparency wavelength $\lambda_{tr-d}$ when the SOA is driven at 70 mA, so the transparency wavelength $\lambda_{tr-m}$, as shown in Figure 3.5, is close to 1479 nm as well. Thus it is possible to shift the $\lambda_{tr-m}$ to wavelengths longer or shorter than 1479 nm by tuning the SOAs driving current, in order to perform either gain depletion or gain pumping at 1479 nm.

To investigate the pumping characteristics around 1479 nm and its relation with driving current and carrier density, the 1479-nm pump light is launched together with a weak continuous wave (CW) probe light at 1547 nm into the SOA in a counter-
propagating direction [47]. We sweep the driving current from 25 mA to 100 mA and the pump light power from 0 to 100 mW, and the measured output power of the probe light is shown in Figure 3.7. When the driving current is about 58 mA, the output probe power stays nearly the same as the input, independent of the increasing pump light power. Therefore the material transparency wavelength ($\lambda_{tr-m}$) is approximately 1479 nm when the SOA is driven at 58 mA. In the relatively low (less than 58 mA) current region, the probe light experiences more gain as the pump light power increases, which indicates that the SOA is absorbing energy from the pump light to increase the carrier density and boost the probe light. Thus, under low driving current (less than 58 mA) conditions, 1479 nm is still in the absorption region of the gain curve. By contrast, when the SOA is driven above 58 mA, the output power of the probe light drops dramatically as the incident power of 1479-nm pump light increases. This occurs because with high driving current, the gain curve is shifted to a shorter wavelength, such that the 1479-nm pump falls into the gain region and depletes the gain. Consequently the pump light is switched from gain pumping to

Figure 3.6: Fiber-to-fiber gain/loss at 1479 nm as a function of driving current.
gain depletion when the driving current increases beyond 58 mA. The pump light can be configured to modulate the probe light intensity either in an inverted way (gain depletion) or non-inverted way (gain pumping). However, it can be observed from Figure 3.7 that the gain of the SOA cannot be constantly increased by optical pumping. As the pump power rises, the output probe power changes less. This occurs because an extremely strong pump increases the carrier density and pushes the gain curve towards the transparency wavelength, where the absorption efficiency \(a_d\) becomes so small that the SOA can hardly absorb additional energy from the pump light; in other words the SOA becomes saturated. For gain depletion, the output probe power has a similar behavior. Therefore all the curves tend to converge as the pump power increases, as shown in Figure 3.7, where the converged value is defined by the output power when the SOA is operating at the transparency point.

![Figure 3.7: Gain pumping/depletion effect in the SOA induced by a 1479-nm pump for different SOA driving currents.](image)

Figure 3.7: Gain pumping/depletion effect in the SOA induced by a 1479-nm pump for different SOA driving currents.
To study the efficiency of the cross-gain modulation by the 1479-nm pump light, we examine the extinction ratio (ER) by switching the pump on and off \[47\]. Figure 3.8 illustrates that the ER varies as a function of driving current. Here, negative ER represents inverted output in polarity with respect to the input (gain depletion). For the gain pumping effect (ER > 0), the ER increases as the driving current is reduced, corresponding to higher modulation efficiency. This occurs because lower carrier density shifts the gain curve to longer wavelengths, and pushes the pump deeper into the absorption region. Thus the pump should have a larger $a_d$ and larger power margin before the SOA saturates.

![Figure 3.8: On-off extinction ratio as a function of SOA driving current.](image)

With both gain pumping and depletion effects available, an SOA is able to behave as a versatile integrator taking both excitatory and inhibitory inputs, and generate non-inverted integration result. To explore this possibility, we investigate and demonstrate the simultaneous operation of gain pumping and gain depletion effects in a single SOA, and that the two opposite effects can cancel each other by proper
control of the operating point. In this experiment, two optical pump signals at 1479 nm (for gain pumping) and 1547 nm (for gain depletion) co-propagate in the C-band SOA driven at 34 mA, while a probe light at 1553 nm is injected in the opposite direction. Both pump signals are modulated with a 10-Gb/s on-off keying pattern ‘10001010’. Figure 3.9(a) illustrates the input pattern of the 1479-nm pump, where 1547-nm pump has the same pattern but lower power. We first switch on only one of the optical pumps. As shown in Figure 3.9(b), when the 1479-nm signal (gain pumping) is on, the output probe signal is a non-inverted copy of the input pattern. On the other hand, with only the 1547-nm signal (gain depletion) present, the output is inverted with respect to the input, as in Figure 3.9(c). It is worth noting that since the driving current is relatively low at only 34 mA, the cross-gain modulation result for either gain pumping [Figure 3.9(b)] or gain depletion [Figure 3.9(c)] shows a slow gain-recovery process. When both of them are on, as in Figure 3.9(d), the output is a straight line with no data. This indicates that the carrier density changes due to gain pumping and gain depletion effects balancing each other, and resulting in no change.

However, to act as an integrator the performance shown in Figure 3.9 is not good enough. Although in Figure 3.9(b) and (c) a temporal integration can be seen, especially for the last four digits ‘1010’, the magnitude difference of the output from the SOA is too small for the subsequent thresholder to distinguish whether the results is above the threshold or not. There are four main reasons resulting in the small magnitude difference. Firstly, since the pump light to boost the carrier density is at 1479 nm, which is quite close to the SOA’s transparency wavelength, the SOA has to be driven at a fairly small driving current (34 mA) to make sure that 1479 nm still lies in the gain pumping region. Such a small driving current makes electrical pumping (resting carrier density \(N_{\text{rest}}'\)) quite low, resulting in a very low output power level from the SOA, as shown in Figure 3.9 (b), (c) and (d). Therefore, although we can clearly see the integration result varies with input in Figure 3.9(b) and (c), the
Figure 3.9: Experimental results of simultaneous operation of gain pumping and depletion: (a) Input pattern of 1470-nm pump. Cross-gain modulation (b) with only gain pumping, (c) with only gain depletion and (d) with both.

absolute output power level is fairly small. Secondly, due to the fact that 1479 nm is very close to the transparency wavelength, we have to keep the power of input optical pulses low enough [Figure 3.9(a)] to prevent the SOA from becoming saturated. Also because it is too close to the transparency wavelength, the 1479-nm light has a low efficiency at pumping the gain of the SOA, as illustrated in Figure 3.7. As a result, the relative power changes between the peaks and valleys are too small for the thresher to discriminated. Thirdly, the 1479-nm CW light is carved into pulses in Figure 3.9(a) by a C-band EOM, whose performance is not as good as in the S-band. It can be seen in Figure 3.9(a) that after pulse curving, the extinction ratio is only 8 dB and at the zero level there is still 100-µW light power injected into SOA, which shrinks the magnitude difference of the output even more. Lastly, as mentioned earlier the 1479-nm laser is intended for use as a pump source in EDFAs, so its output has a relatively poor quality in terms of intensity noise, implicated by the
thick ‘1’ level and ‘0’ level in Figure 3.9(a). Consequently the output has a relatively large intensity noise as well, which makes it even more difficult for the thresher to distinguish ‘1’s and ‘0’s. Thus due to the lack of a high-quality laser source and a proper pulse-curving modulator at shorter wavelengths, the performance of the SOA as an integrator is greatly impaired. However, customized components specially designed for shorter wavelengths or an SOA with a longer transparency wavelength might make this versatile integrator practical.

Other than the SOA, another electro-optic device—the electro-absorption modulator (EAM) has a slow recovery characteristic of its carrier density, rendering it suitable for temporal integration as well [48]. Its cross-absorption modulation (XAM) [55], opposite to XGM in SOA, functions as an inverted gain pumping effect. Based on Franz-Keldysh effect and quantum-confined stark effect, the absorption in the EAM is larger under a negative voltage bias, such that the input optical signal is absorbed and generates carriers. The optically generated carriers screen the external electric field as well as cause the band filling and exciton bleaching effect, which results in shifting the absorption spectrum edge to the shorter wavelength side [56]. The shift in spectrum allows the spike source to pass through the EAM even under strong negative voltage bias, i.e. modulating the intensity of the spike source. The absorption dynamics of an EAM are governed by the rate equation as follows [57]:

\[
\frac{dN(t)}{dt} = \frac{\eta P}{h\nu} - \frac{N(t)}{\tau} \tag{3.4}
\]

where \( P \) is the power of the incident light, \( \eta \) is the ratio of absorbed optical power to total incident power, \( h\nu \) is photon energy, and \( \tau \) is carrier escape time constant which depends on the reverse voltage and carrier density.

Since the carrier escape time \( \tau \) is finite, the carrier density and transmittance decreases gradually over time. This happens shortly after the input signal exited
the EAM. Therefore, any input signal that arrives within the EAM recovery window will be integrated. The output intensity of a probe pulse train reflects the strength of the temporally-integrated input signals. We first measure the recovery time to explore the possibility of using an EAM as temporal integrator. A strong optical pump at 4.6 dBm is used, while a weaker probe pulse is temporally scanned from about 50 ps before the launching of the pump pulse to 550 ps after launching to sample the effect. Both the pump and probe pulses have pulse widths of ∼5 ps.

Figure 3.10(a) shows the measured recovery time of an EAM at different voltage bias points, indicating the larger the bias (more negative), the shorter the recovery time. Due to the finite recovery time of the EAM, the transmittance of the EAM gradually decreases over time, after the influence of an optical pulse. Therefore, by placing multiple control pulses within the recovery interval, the EAM integrates them. Strength of the integrated output depends on the temporal separation of the input signals. The recovery time of the EAM also depends on the optical pump power. Figure 3.10(b) shows the change in recovery time as a function of the optical pump power in the form of optical pump-pulse attenuation.

![Figure 3.10](image)

Figure 3.10: (a) Measured absorption recovery time as a function of EAM bias voltage. (b) Measured absorption recovery time as a function of optical pump-pulse attenuation.

To study the integration effect in an EAM, two strong input optical pulses of about the same power (4 dBm) are used, while a weaker optical probe pulse at −2
dBm is used to measure the effect. The separation of the two pump pulses is variable, while the probe signal is temporally scanned from about 50 ps before the launching of the first pump pulse to 550 ps after the launch. Shown in Figure 3.11 is the output probe signal power (i.e. the transmittance of the EAM) after integration, when the EAM is biased at $-2.5 \text{ V}$. The black data points represent the output with just one input pulse. When the second pump pulse is placed 100 ps after the first pump pulse, no integration is observed, as shown by red data points. This occurs because the second pump pulse is outside the integration window, i.e., the EAM completely recovers before the second pump pulse is launched into the EAM. As the second pump pulse moves closer to the first pump pulse, time-sensitive integration is observed, resulting in a different integrated power depending on the temporal separation between the two pump pulses. As seen from the purple data points, where the temporal separation between the two pumps is 20 ps, the power is much higher than for other separations because the output of the time-sensitive signal integrator now represents an integration of two closely-spaced pulses.

After studying the integration behavior of the EAM-based temporal integrator, a more complex signal is used to verify the performance of temporal integration in a photonic neuron, as shown in Figure 3.12(a), measured by a photoreceiver at 30 GHz. To measure the dynamic change in the EAM, we use the same probe scanning approach as describe above. The signal consists of pulses of different heights and temporal separations. There are several features worth noticing, as indicated by (i)–(iv). The integrated output with a long integration time (EAM biased at $-1 \text{ V}$) and short integration time (EAM biased at $-2.5 \text{ V}$) are shown in Figure 3.12(b) and (c), respectively. The red dotted lines are used as a reference to compare the integration behaviors of the two cases. Feature (i) consists of two pulses that are very close together (<10 ps). The integrated output in both the $-1$-V and $-2.5$-V bias cases shows a strong integrated signal result. For (ii), the three pulses are of medium
Figure 3.11: Integration response of the EAM with different temporal separation of the pump pulses.

separation, such that they can only be integrated and result in a strong intensity if the integrator has a longer integration window, as shown in Figure 3.12(b). For (iii), the first two pulses are very close together, while the third pulse is further apart. Thus, the first and second pulses are integrated in both Figure 3.12(b) and (c) and give a relatively strong output (above the red dotted reference line). In the −1-V bias case, due to the longer integration time, the third pulse also is integrated and results in a strong output. For (iv), the three pulses are further apart and none of the −1-V and −2.5-V bias cases has a sufficiently long integration time to obtain significant integration. Therefore, the output pulses are below the red reference line for both cases. From Figure 3.12, we experimentally study the dynamic of signal integration in a photonic neuron, and the integration can be configured to have different integration responses towards the input signal.
Figure 3.12: Time sensitive signal integration in EAM: (a) input signal, (b) integrated output with EAM biased at $-1 \text{ V}$ (long integration window), and (c) integrated output with EAM biased at $-2.5 \text{ V}$ (short integration window).
3.2.2 Enhanced thresholder

Besides the integrator, the optical thresholder is another critical element in a photonic neuron. As discussed in Section 3.1.3, a Ge-doped NOLM (Figure 3.3) is employed to implement all-optical thresholding. Its transfer function is plotted in Figure 3.4 where the flat top at normalized input power set to 1 is used to unify the magnitude of output spikes above threshold, while the flat bottom at around zero input power is used to eliminate the small spikes bellow threshold. It was observed that flatness and width of the two regions determine the output quality and input power fluctuation tolerance. If the input power is too high, far beyond 1 and reaching the valley at around 1.5, then the thresholder will induce pulse bifurcation to the output pulse, because the input pulse’s rising and falling edges cross the top of the transfer function twice in total. In addition, since the transition between 0 and 1 of the normalized input power curve is not a perfect step function, it is difficult to do thresholding of the pulses located in the middle of the transition region. Moreover, the further from 0 or 1 the pulses are located, the more noisy the output from the thresholder is. Thus to improve the thresholder’s performance, it is critical to build flatter and wider top and bottom regions and a steeper transition region in between. Unfortunately the current version of thresholder—Ge-doped NOLM does not offer the tunability of these parameters.

To build a more powerful threholder, my labmate Nicole Rafidi and I propose and demonstrate tailoring of the power transfer function of the Ge-doped NOLM with a low requirement on the nonlinear phase shift using offset-spectral filtering [50]. Power transfer function tailoring makes the thresholder more flexible and allows customization of the thresholding parameters to the requirements of various systems. Particularly in our photonic neuron, it has the capability to adjust the power transfer function and the thresholding slope and to reduce pulse distortion due to pulse bifurcation.
Figure 3.13 shows the architecture of the proposed scheme. The modified NOLM consists of an asymmetric coupler (90:10), 15-m of highly Ge-doped nonlinear fiber (HDF) (preform 311cf) [58], a tunable isolator, and a polarization controller. At the output of the modified NOLM, a tunable optical band-pass filter with a 3-dB bandwidth of 0.38 nm is used for offset-spectral filtering. The optical band-pass filter is a three-stage Gaussian filter, where a filtering profile with steeper cutoff is expected to give a larger improvement in the power transfer function. As an input pulse source, we use a 1.25-GHz repetition rate mode-locked fiber laser generating ~3-ps width pulses at the wavelength of 1551 nm. The laser output is amplified by an EDFA to ensure nonlinearity is induced in the Ge-doped nonlinear fiber. Purple arrows correspond to the clockwise beam, while the orange arrows correspond to the counterclockwise beam.

Figure 3.13: Schematic of enhanced NOLM with BPF. BPF: Optical band-pass filter, PC: polarization controller, ISO: tunable isolator; HDF: highly Ge-doped fiber (preform 311cf). Thicker arrows correspond to a larger power signal.
The Ge-doped NOLM is a Sagnac interferometer that measures the differential phase shift of the input pulse caused by the nonlinear phase shift in the HDF due to asymmetric splitting of the input. The input pulse is split into two counter-propagating pulses, one with 90% of the original power, and one with 10% of the original power. The size of the arrows in Figure 3.13 represents the amplitude of the pulse the arrow describes. When each pulse reaches the nonlinear fiber, it incurs a phase shift proportional to its power. Thus, the stronger pulse has a significantly greater phase shift than the weaker pulse. The tunable isolator acts as a directional attenuator, allowing the weak pulse to pass and attenuating the counter-propagating strong pulse. The tunable isolator is tuned to balance the power of the counter-propagating pulses at the coupler output. Thus, when the two pulses meet at the coupler after completing the loop, they have the same power but different phases. The polarization controller allows for fine adjustment of the relative phase of the pulses, according to the Jones matrix. The interference is such that for low powers, they interfere destructively at the output, while for high powers (above the threshold level), they interfere constructively at the output. Due to SPM in the HDF (preform 311cf, doped with 75 mol.% GeO$_2$) with a large nonlinear coefficient of 35 W$^{-1}$km$^{-1}$, the optical spectrum at the output of the modified NOLM is broadened. The output is then passed to a tunable optical band pass filter for offset-spectral filtering. By off-setting the filter from the center frequency of the pulse, the pulses that are above threshold are selectively passed through the filter, as determined by the filter position, improving the thresholding capability of the device. Although a change in wavelength results after offset spectral filtering, the original wavelength can be restored through wavelength reconversion techniques.

To examine the improvement of the Ge-doped NOLM with a BPF compared to the Ge-doped NOLM alone, the bandpass filter’s center wavelength is adjusted to achieve a step response for thresholding. The location of the filter center wavelength is
Figure 3.14: Power transfer function, with pertinent regions labeled. This transfer function was taken for the Ge-doped NOLM device. The ideal power transfer function resembles a Heaviside step function, with 0 slope in the high and low level regions, and infinite slope in the linear region.

extremely important in determining the efficacy of the thresholding. This determines the steepness of the slope of the resulting power transfer function, the flatness of high level transmission, and efficiency.

In order to characterize the performance of the Ge-doped NOLM with BPF as a thresholder, we plot the power transfer function, i.e., the output power versus input power, normalized in both x- and y-axes to the occurrence of the step (or threshold value). The parameters for characterizing the performance of the thresholder are (i) polynomial order of the linear region in the power transfer function and (ii) flatness of the low and high output power levels. Figure 3.14 illustrates these regions. In order to detect these regions we find the localized slopes of the power transfer function. When the slopes approach a constant value, this determines the slope of the linear region. The region of higher power after the linear region is the high-level region.
The power transfer functions are normalized to the start of this region. Polynomial order is measured by fitting a straight line to the linear region on a log-log plot. The slope of the fitted line is the polynomial order of the device. A high polynomial order is desirable, because it results in better ability to distinguish between low and high signal values. This is balanced against the desire to have a stable high-level region after the threshold, and to suppress one-level noise.

Figure 3.15: Power transfer functions of Ge-doped NOLM and various offsets of the BPF. The numbers in the legend correspond with the offset of the center wavelength in nm of the BPF from 1551 nm for that data set.

In our experiment, we performed a sweep using pulses of different input powers to get the power transfer function of our Ge-doped NOLM with BPF device and compare it to that of the Ge-doped NOLM alone, which we also obtained experimentally. The results are shown in Figure 3.15. It is seen that the high level of the Ge-doped NOLM transfer function has a large variation in power, which limits the ability of the Ge-doped NOLM to suppress high-level noise. As the center wavelength of the bandpass filter moves further away from the input signal wavelength, a more stable
Figure 3.16: Low-level region of the power transfer function, showing that a better low-level performance is obtained with the Ge-doped NOLM with BPF.

high-level region results, which indicates that a better high-level noise suppression can be obtained. The experimental results indicate that a spectral offset between 2.5 nm to 2.75 nm yields the most stable high-level region. As for the low-level suppression, Figure 3.16, which is an expansion of the low-level region of Figure 3.15, shows that the use of BPF greatly enhances the performance of the Ge-doped NOLM, resulting in a flattened low-level region.

The thresholding ability of a given technology can be better quantified by examining the polynomial order of the linear region of the power transfer function, i.e., steepness of the slope, as shown in Figure 3.17. Here, a steeper slope is representative of better thresholding ability. The slope is taken from the curves shown in Figure 3.15. As shown in Figure 3.17, adding a BPF to the Ge-doped NOLM makes the slope of the thresholder much steeper, and a larger offset in spectral filtering results in a steeper slope in the transfer function, indicating better thresholding ability despite the decrease in power efficiency of the device. We have simulated and studied
how the change in BPF bandwidth affects the power transfer function. Despite the
decrease in output power, we observed that a narrower bandwidth results in a flatter
low level region as well as a steeper slope, while the flatness of the high-level region
evolves as the bandwidth of the filter changes. The evolution of high-level region flat-
ess greatly depends on the input pulse profile, filter shape, as well as the wavelength
offset of the bandpass filter.

![Diagram](image)

Figure 3.17: Polynomial order, the slope taken from a linear fit to a log-log plot of
the linear region of Figure 3.16. The horizontal line marks the order of the power
transfer function of the Ge-doped NOLM.

We have also studied how the use of BPF improves the thresholding capability
of Ge-doped NOLM by experimentally examining its effect on a noisy eye diagram.
This helps put Figure 3.15 and Figure 3.17 in a different perspective, showing the
overall effect of the shape of the power transfer function on the signal-to-noise ratio.
Figure 3.18 shows the eye diagrams of (a) a noisy input, (b) output after the Ge-doped
NOLM, and (c) output after the Ge-doped NOLM with BPF, captured by a 30-
GHz oscilloscope. Figure 3.18(a) shows the input signal with pseudo-random binary sequence (PRBS) pattern, which has been corrupted by improperly biasing the EOM to introduce noise to the signal and an undesired level at the zero-level. Figure 3.18(b) shows the signal after the Ge-doped NOLM, where moderate improvement is obtained with residual noise at both high and low levels. Figure 3.18(c) shows the pulse after the Ge-doped NOLM with BPF, which has greatly reduced noise, flattening both the bit 1 and bit 0 levels. It demonstrates that the use of offset-spectral filtering by an optical BPF significantly improves the performance of the Ge-doped NOLM.

Figure 3.18: Eye diagram comparison. (a) The input signal to the system. Noise was added before the Ge-doped NOLM. (b) Signal after the Ge-doped NOLM. (c) Signal after Ge-doped NOLM with BPF configuration.

Figure 3.19 shows the results for improvement in pulse bifurcation by applying offset spectral filtering after the Ge-doped NOLM. The simulation model included a Ge-doped NOLM with a 0.38-nm bandwidth bandpass filter, offset by 2 nm from the central input wavelength. In the simulation, a 5-ps wide pulse is used as the input. Offsetting the filter from the central wavelength corrects pulse bifurcation and thus improves overall device performance.

In the experiments, we were unable to measure pulse profiles directly due to the limited bandwidth of the oscilloscope (see Figure 3.19). As shown in Figure 3.19(a), pulse bifurcation does not occur in the low power configuration; however, as the power increase, the pulse is split into two parts, as shown by the red square data points. With the use of offset-spectral filtering, no pulse bifurcation is observed, as
shown in Figure 3.19(b). However, one trade-off is that the pulse width is significantly broadened after filtering due to the narrower passband of the BPF compared to the spectral width of the input pulse. Therefore, input pulses were not perfectly restored at the output. However, this experimental limitation does not diminish our results, as the improvement we can get from the Ge-doped NOLM with BPF configuration is more significant. The pulse width can be restored through post-compression of the output pulse.

![Figure 3.19: Simulation results for improvement in pulse bifurcation using Ge-doped NOLM with BPF. Three pulses are in the low, linear, and high region of the power transfer function in accordance with amplitude. (a) Pulse shapes outputted by the Ge-doped NOLM. (b) Pulse shapes output by the Ge-doped NOLM with BPF.](image)

### 3.2.3 Asynchronous spiking photonic neuron

The first photonic neuron presented earlier in Section 3.1 realized most features of spiking processing, as we discussed in Section 2.3. However it still lacked one critical feature—asynchronous behavior. In spike processing, the information is carried by analog timing position instead of digital timing position that is synchronized to a clock. This analog timing makes sure that spike processing keeps high efficiency of analog systems. However in the first photonic neuron, a sampling pulse train is needed as a spike source to present the integration result in the temporal integration-
tor. In other words, digitized timing position with fairly high resolution is used to emulate the truly analog spiking behavior. Consequently, although input spikes are asynchronous in time, the output spiking time is dependent on and synchronized with the sampling pulse train, which makes the photonic neural response synchronous and induces quantization error in spike timing. As a result, the capacity of the neural output is limited by the sampling rate, which harms the efficiency of the photonic neuron. Higher sampling rates can reduce quantization error, but it might cause multiple spikes triggered by a single stimulation.

To overcome this issue, we propose and demonstrate a photonic spiking neuron with truly asynchronous response to stimulations [48][49]. In the asynchronous photonic neuron, a spike source is still needed to provide spikes for the neuron as output. However, the spike source is not a pulse train as in Section 2.3. It is worth noting that a spiking neuron spikes only if the temporally integrated input signals are strong enough to trigger the threshold, otherwise, the neuron will not spike. Therefore, the spike source needs to be present at the integrator only if there is an input signal present at that particular temporal position. Instead of using a pulse train with a fixed repetition rate, a spike source in our asynchronous spiking photonic neuron is generated by making a copy of the input signal at another wavelength through four-wave mixing (FWM) [39], allowing the photonic neuron to spike asynchronously.

Figure 3.20 illustrates the proposed asynchronous spiking photonic neuron. In step (i), input spikes, i.e. Input 1–Input N, are weighted and delayed before they are combined by a coupler, and then split into two portions with exactly the same pattern and same wavelength. One portion is directed to the integrator as inputs [step (ii)], while the other portion is converted into another wavelength by FWM [step (iii)] and then sent to the integrator as a spike source. The inputs and spike source are precisely aligned in time. In the EAM-based integrator (see Section 3.2.1), the strong input spikes deplete the absorption which applies a time-varying magnitude onto the weak
spike source, labeled as step (iv). Subsequently in step (v), the integration result presented by the spike source is measured and processed by the thresholder and then released to the output [step (vi)].

We performed an experiment to demonstrate the basic function of spike processing to verify this asynchronous photonic neuron. Figure 3.21 shows the experimental setup of the asynchronous photonic neuron. Two input pulses of \(~9\)-ps width, both at 1550.12 nm, are weighted and delayed, before they are combined by a 50/50 coupler C1. One output branch of C1, as a probe signal, is launched together with strong CW pump light at 1547.21 nm into 60-m photonic crystal fiber for FWM. The average power of the probe signal and pump signal are 14 dBm and 21.6 dBm, respectively. After the photonic crystal fiber, a 0.5-nm-wide optical filter is used to pick out the newly generated 1544.52-nm signal as the spike source or sampling pulses for the integrator. The input pulses from the other output branch (point A) of C1 are aligned 5ps ahead of the corresponding sampling pulses (point B) in time. After optical amplification, the input and the spiking source are launched into an EAM (CIP 10G-
PS-EAM-1550) for integration, where the integration window is set to 50 ps by fine-
tuning the reverse bias voltage of the EAM. Another optical filter centered at 1544.52
nm selects the spikes output from the EAM integrator. These integrated pulses are
then amplified by an EDFA to compensate optical power losses, and filtered by a 0.5-
nm-wide filter to remove out-of-band amplified spontaneous emission (ASE) noise.
At the end, a NOLM based on 20-meter Ge-doped fiber thresholds the integrated
pulses and produces output spikes (point D).

Figure 3.21: Experimental setup of an asynchronous spiking photonic neuron. W:
weight; t: time delay; Att: attenuator; C1, C2, C3: optical coupler; HNF: highly
nonlinear fiber.

The performance of the photonic neuron is studied in three cases, where the time
interval between input pulses from different branches and weight of the inputs are
different, as shown in Figure 3.22. The configurations Case I: two input signals are
35 ps apart with average power of 7.8 dBm; Case II: two input signals are 85 ps apart
with average power of 7.8 dBm; and Case III: two input signals are 200 ps apart with
average power of 9.8 dBm. The oscilloscope traces of the inputs for all three cases
are shown in Figure 3.22(a)–(c), respectively. Under FWM, the spike source for each
case is generated and illustrated in Figure 3.22(d)–(f), they all have the same pattern
as the corresponding input signals but a different wavelength. In case I (left column
in Figure 3.22), the two input signals have a temporal separation that is within
the integration window of the photonic neuron. Due to the low input signal power, the
first input pulse only saturates the absorption of the EAM slightly, such that only
a small portion of the first sampling pulse can pass through the EAM. Since the
Figure 3.22: Experimental results of input, spike source and output pulses in three different cases with different time intervals and weight values. The three cases we studied are Case I: two input signals are 35 ps apart with average power of 7.8 dBm; Case II: two input signals are 85 ps apart with average power of 7.8 dBm; Case III: two input signals are 200 ps apart with average power of 9.8 dBm.

A second input pulse arrives before the EAM recovers and further saturates the EAM, a larger transmittance of the second pulse in the spike source is observed, as shown in Figure 3.22(g). The output spike is then passed to the Ge-doped NOLM thresholder, such that the first weak output spike is removed and amplitude noise of the second output spike is suppressed. As shown in Figure 3.22(j), the final output consists of only one spike that is triggered immediately after the second input spike. In case II, (middle column in Figure 3.22), input pulses are the same as in Case I, but are now further apart from each other, i.e. outside of the integration window. After the
absorption in the EAM is slightly saturated by the first pulse, the EAM recovers back to its high absorption level before stimulation from the second pulse. Thus, the second pulse induces a similar amount of absorption as the first pulse, resulting in two weak spikes at the EAM output, as shown in Figure 3.22(h). Since both the weak output spikes are below the threshold of the thresholder, both of them are eliminated by the thresholder, as shown in Figure 3.22(k). In case III (right column in Figure 3.22), the two input pulses are much further apart, but have a larger weight. Both the first and second input pulses strongly saturate the absorption of the EAM, resulting in two strong output spikes from the photonic neuron. The two output spikes are above the threshold level of the thresholder and both of them emerge from the thresholder, as shown in Figure 3.22(l). In conclusion, a fully asynchronous spiking photonic neuron is demonstrated.

3.3 Lightwave Neuromorphic Circuits

Neuromorphic engineering provides a wide range of practical computing and signal processing tools by exploiting the biophysics of neuronal computation algorithms. Existing technologies include analog very-large-scale integration (VLSI) front-end sensor circuits that replicate the capabilities of the retina and the cochlea [59]. The spiking photonic neuron can be utilized to provide high-speed and low latency performance to meet the requirements of real-time signal processing. Based on the bench-top prototype of the photonic neuron detailed in the previous sections, here we present a couple of prototypical lightwave neuromorphic circuits, including signal feature recognition [12] and a timing-jitter insensitive logic gate [13].
3.3.1 Signal feature recognition based on crayfish tail-flip escape response

Crayfish escape from danger by means of rapid escape response behavior. The corresponding neural circuit is configured to respond to sudden stimuli. Since this corresponds to a life-or-death decision for the crayfish, it must be executed quickly and accurately. A potential application of the escape response circuit based on light-wave neuromorphic signal processing could be for pilot ejection from military aircraft. Based on the bench-top photonic neuron prototype introduced in Section 3.1 and Section 3.2, we propose and build an all-optical device for signal feature recognition, which mimics the crayfish circuit using photonic technology, and is sufficiently fast to be applied to defense applications in which critical decisions need to be made quickly while minimizing the probability of false alarm [12].

Figure 3.23 illustrates the crayfish escape neuron model. Signals from the receptors (R) are directed to the first stage of neurons—the sensory interneuron (SI), which is configured to respond to specific stimuli at the receptors. The SI integrates the stimuli and generate spikes when the inputs match the default feature. The spikes are then launched into the second stage of the neural circuit—the lateral giant (LG). The LG integrates the spikes from the first stage and one of the receptor signals. The neuron responds only when the signals are sufficiently close temporally and strong enough to trigger a spike.

The signal feature recognition circuit consists of two cascaded EAM-based integrators (see Section 3.2.1) and one optical thresholder (see Section 3.1.3). The first integrator is configured to respond to a set of signals with specific features, while the second integrator further selects a subset of the signal from a set determined by a weighting and delay configuration and responds only when the input spike and the spike from the first integrator arrive within a very short time interval.
In the experimental setup shown in Figure 3.24, inputs $a$, $b$ and $c$ are weighted and delayed such that the first EAM (CIP 10G-PS-EAM-1550) integrator (EAM 1) is configured to spike for inputs with specific features. A train of sampling pulses is launched together with the inputs to provide a pulsed source for the EAM to spike. The spiking behavior is based on XAM [55] in an EAM, as demonstrated in Section 3.2.1. The integration window—recovery time of the EAM is depicted in the inset in Figure 3.24. The spike output is then thresholded at the NOLM-based thresholder, as presented in Section 3.1.3. The thresholded output and part of input $b$ are launched into the second integrator (CIP 10G-LR-EAM-1550) (EAM 2) as the input control through path $\beta$ and $\alpha$, respectively. Sampling pulses are launched to the integrator through

---

**Figure 3.23:** Schematic illustration of the crayfish tail-flip escape response.

**Figure 3.24:** Schematic illustration of the optical implementation of the escape response. $W$, weight; $t$, delay; EAM, electro-absorption modulator; TH, optical thresholder. Inset, measured recovery temporal profile of cross-absorption modulation in the EAM.
path $\xi$ as a spiking source. Through weighting and delaying of the inputs, spikes by the second integrator occur only for inputs with the desired features. The selection of the desired features can be reconfigured simply by adjusting the weights and delays of the inputs. Figure 3.25(a) shows the input control signal consisting of one, two, or three pulses that are within the integration window. After integration at the EAM, the sampling pulses representing the integrated output are shown in Figure 3.25(b) with the superimposed temporal profiles shown in the insets. The EAM is adjusted to spike when the input has two or more pulses occurring within the integration time window. To equalize the heights of the output and remove the unwanted weak pulses, the integrated output is sent to the following optical thresholder.

![Figure 3.25: (a) Input to EAM. (b) Output sampling pulses of EAM. Insets, superimposed output temporal profiles.](image)

In this signal feature recognition experiment, there are three inputs, $a$, $b$, and $c$. The recognition circuit detects input patterns of $abc$ and $ab$- having specific time intervals between the inputs, as shown in Figure 3.26. Figure 3.26, part (i), shows five different input combinations, where input $a$ is designed to have the largest weight, while input $b$ has the smallest weight. The separation of each set of inputs is sufficiently large that they do not affect each other’s integration performance. Five
Figure 3.26: Schematic illustration of configuring pattern recognition. (i)–(iv) First integrator and thresher. (v)–(vii) Setting 1, detection of $abc$ and $ab\bar{c}$. (viii)–(x) Setting 2, detection of $abc$ only. (xi)–(xiii) Setting 3, exceed integration window–no input is detected.
sampling pulses are used for each set of inputs, as shown in Figure 3.26 part (ii). With the use of multiple sampling pulses, high temporal resolution of integration, as well as differing spike patterns for the second stage of integration, can be obtained. The input pulses change the transmittance of the EAM 1 as shown by the profile curves in Figure 3.26 part (iii); thus the amplitude of the sampling pulses (shaded pulses) at the EAM 1 output also changed. The output spikes from EAM 1 are then thresholded at the level indicated by the dotted line. The spikes are now of similar amplitude and the undesired weak spikes are removed [Figure 3.26 part (iv)]. Different spike patterns are obtained for different inputs. In this experiment, we identify patterns \( abc \) and \( ab \). Here, we use the EAM 1 output as the sampling pulse, which also helps eliminate undesired patterns; i.e., input \( a \) is absent. The EAM 1 output is launched into the second integrator as sampling pulses, while a duplicated, delayed, and weighted copy is combined with input \( b \) and used as the control for the second integrator [Figure 3.26 part (v)]. By adjusting the relative temporal delay between the input \( b \) and the inputs to the EAM 2, different input features are recognized, as shown in Figure 3.26 parts (v)-(xiii).

Figure 3.27 shows the experimental measurements of the signal feature recognition. In the experiment, the input signal has an average power of \( \sim 2.5 \) dBm. Figure 3.27(a) shows all eight combinations of the three inputs with specific weights and delays. We use ‘1’ to represent the presence of input, while ‘0’ means there is no input. A superimposed temporal profile of the input signal is shown in the inset of Figure 3.27(a). Sampling pulses with power of 9 dBm and separation of \( \sim 25 \) ps are used, as indicated by the arrows. The input signals are integrated, and the transmittance of EAM 1 is represented by the spike pattern at the output [Figure 3.27(b)]. A Ge-NOLM is used to threshold the output of EAM 1 [Figure 3.27(c)]. The inset shows the superimposed temporal profile of the thresholded output. FWM in a 35-cm bismuth-oxide nonlinear fiber is used [60] to duplicate the thresholded output. Two copies at different wave-
lengths are obtained, one as the sampling pulses and the other one as the input to the EAM 2. The sampling pulses have optical power of 1.25 dBm, while the input signal to the EAM 2 is at 4.7 dBm. When the output spikes from the first neuron arrive at the second integrator slightly after input $b$, i.e., within the integration interval, the second integrator will spike. By adjusting the time delay of the inputs to the 2, the pattern recognizer identifies patterns $abc$ and $ab$- [Figure 3.27(d)] or just $abc$ [Figure 3.27(e)]. However, when the spikes from the first neuron arrive too late, i.e., exceed the integration interval, the second integrator will not spike [Figure 3.27(f)].
These examples indicate that the signal feature recognizer performs correctly and is reconfigurable through time-delay adjustment.

3.3.2 Timing jitter insensitive photonic logic gate

As the speed of telecommunication and computing dramatically increases and reaches the physical limit of electrical devices, all-optical logic gates have received considerable attention in the field of next-generation optical networks and optical computing systems for optical signal processing functions [61][62][63][64][65][66][67][68]. Several approaches have been demonstrated to realize various logic functions using highly nonlinear fibers [61] or semiconductor optical amplifiers (SOA) [62][63][64][66][67]. SOA-based devices are promising due to the advantages of being compact, highly power efficient, and having the potential of monolithic integration. So far, most logic gates based on SOAs have been implemented by employing XGM [62][63][64][65], cross-phase modulation (XPM) [66], FWM [63][65][67] and nonlinear polarization rotation [68][69].

In digital communication systems and networks, timing jitter is some deviation, usually undesired, of the temporal signal positions from perfect periodicity of an assumed periodic signal. The smaller timing margins associated with today’s high-speed communication and signal processing systems reveal that tighter control of jitter is critically needed throughout the system design. For photonic logic gates, which operate at gigahertz level and deal with picosecond optical pulses, timing jitter imposes even more critical limits. Unfortunately none of the existing all-optical logic gates is immune to the timing jitter of the input signals.

Based on the photonic neuron prototype, we design and demonstrate a timing jitter insensitive photonic logic gate device, which can be reconfigured between AND and NOT functions [13]. It employs the SOA-based versatile integrator (see Section 3.2.1) and the NOLM-based optical thresholder (see Section 3.1.3). Figure 3.28
illustrates the setup of the reconfigurable timing jitter insensitive logic gate. It has
two input ports for pulse inputs, with the driving current and threshold level as its
tunable parameters. The 1479-nm input pulses are received at input port 1 and
launched into the C-band SOA (Kamelian NL-L1-C-FA-1550). At this wavelength,
the optical pulses can either create gain pumping or gain depletion effects in the SOA
depending on the driving condition, as demonstrated in Section 3.2.1. The 1550-nm
pulses launched into input port 2 travel through the SOA in the opposite direction,
and experience the gain change in the SOA. Then the following thresholder based on
a Ge-doped NOLM eliminates the pulses under the threshold level and unifies the
pulse power with higher power level (see Section 3.1.3).

Figure 3.28: Experimental setup of reconfigurable and timing jitter insensitive logic
gate. CW: continuous wave; PC: polarization controller; MZM: Mach-Zehnder mod-
ulator; SOA: semiconductor optical amplifier.

When the SOA is driven by low current, specifically 15 mA in our experiment,
the 1480-nm light generates the gain pumping effect [47]. Thus every 1480-nm pulse
from Input 1 increases the carrier density, according to the rate equation 3.3 and
consequently boosts the gain of the SOA. The counter-propagating 1550-nm pulses
experience the gain pumping, resulting in an amplitude change. If the threshold level
is properly tuned, the thresholder can let through the 1550-nm pulses boosted by
1480-nm pulses but block the ones without the gain pumping from 1480-nm pulses.
Therefore, at the output port, only if 1480-nm and 1550-nm inputs are both present, the output emerges from the logic gate, which corresponds to an AND operation on Input 1 and Input 2.

With high driving current injected into the SOA, 75 mA in this case, the 1480-nm light changes its role to depleting the gain. The 1550-nm pulses from input port 2 experience very little gain if the gain is depleted by 1480-nm light. Otherwise, the 1550-nm pulses are amplified by the SOA with a higher gain. The thresholder is set to the proper level to distinguish the output in these two cases. Thus, if we keep pulse train at input 2 always available as a probe signal, the output is a NOT operation on input 1, as shown in Figure 3.28. As a result, the proposed logic gate can be reconfigured by simply changing the SOA’s driving current and adjusting the threshold level if necessary.

Another feature of the proposed logic gate is its high timing-jitter tolerance. From the rate equation after the optical pump disappears, the carrier density will exponentially decay back to the rest level with time constant of \( \tau \). Therefore, after each stimulation (gain pumping/depletion) by the 1480-nm pulse, there is a small time window before the carrier density recovers to the original rest state. Within this time window, the 1550-nm pulses can still be affected by the gain pumping/depletion effect. Although the gain pumping/depletion becomes gradually weaker with time, resulting in the output amplitude change, the following thresholder can correct the unequal amplitude and keep the same output level for all inputs within the time window. Moreover, by tuning the threshold level, we can choose to invert the output when the gain pumping/depletion is weakened to a certain level, so as to change the effective time window of the logic operation. In this way, we can reconfigure the function and the timing-jitter tolerance window by adjusting the driving current and threshold level.
Figure 3.29: Experimental results of AND gate: (a) Input 1 at 1479 nm; (b) Input 2 at 1549 nm; (c) Output with Input 1 aligned with Input 2; (d) Output with Input 2 lagging input 1 by 50 ps; Output with Input 2 lagging input 1 by 100 ps. Experimental results of NOT gate: (g) Input 1 at 1479 nm; (h) Probe at 1549 nm; (i) Output with Probe aligned with Input 1; (j) Output with Probe lagging Input 1 by 25 ps; (k) Output with Probe lagging Input 1 by 50 ps.
In the experiment, a mode-locked fiber laser at 1549 nm is used to generate a 10-GHz pulse train, while a 1479-nm continuous wave (CW) light is carved into a 10G pulse train by an EOM. Then the 1479-nm and 1549-nm pulse trains are modulated by different data patterns to get different bit combinations as Input 1 and Input 2 for verifying the logical operations. Figure 3.29(a)–(e) illustrate the experimental results of the AND gate, when the driving current of the SOA is set to 15 mA to enable a gain pumping effect of the 1479-nm pulses. Figure 3.29(a) and (b) show the input 1479-nm and 1549-nm pulses, with 300-ps spacing between adjacent bits. Their pulse widths are ∼30 ps and ∼9 ps, respectively. The 1479-nm pulses have wider pulse width because of wide electrical pulses used for pulse carving. The ripples on the zero level are due to imperfect intensity modulation. Since the 1479-nm light goes through intensity modulation twice, we use an S-band SOA to compensate the power losses, which induces some amplitude noise in Figure 3.29(a). The output from the logic gate is shown in Figure 3.29(c). The output pulse shows up only when both inputs are present, which is consistent with AND logic. To verify the timing-jitter insensitivity, we also investigate the situation when the two inputs are not aligned in time. The 1549-nm pulses (Input 2) are shifted continuously to a later time position, with all other parameters unchanged. With 50-ps lag behind Input 1, as illustrated by Figure 3.29(d), the output keeps the same amplitude level as when there is no lag, as shown through comparison with Figure 3.29(c). The timing lag is increased up to 100 ps [Figure 3.29(e)], until the output pulse level starts to decrease and become noisy, which means the output level from the SOA reaches the threshold level. Thus, the AND gate is capable of tolerating up to ±50-ps timing jitter between the two inputs.

For the NOT gate, the driving current of the SOA is changed to 75 mA, to enable the gain depletion effect by the 1479-nm input. Figure 3.29(g) and (h) show the inputs. The threshold level is adjusted accordingly to distinguish the outputs with
and without the gain depleted. Input 2 at 1549 nm is considered as probe light for the NOT gate. As illustrated in Figure 3.29(i), when a pulse enters Input 1, the SOA cannot provide enough gain to boost the probe pulse above the threshold level, so the output shows no pulse. Therefore the function of a NOT gate is demonstrated. Similar to the AND gate, we also test the timing-jitter tolerance of the NOT gate. The probe pulse is shifted to lag Input 1 pulse by up to 50 ps, when the output level reaches the threshold level. Figure 3.29(j) and (k) show the output results of 25-ps and 50-ps lags respectively. No obvious change is observed comparing the results with the output without timing shift shown in Figure 3.29(i). Therefore the NOT gate allows ±25-ps timing jitter between the input and probe. It is worth noting that, when the SOA operates as an AND gate with 75-mA driving current, the stronger electrical pump makes the carrier density level decay faster back to the resting level than the NOT gate condition with only 15-mA driving current. Thus the NOT gate has a narrower timing-jitter tolerant window than the AND gate.

3.4 Conclusion

In this chapter, we presented the elemental design of the first photonic neuron, which emulates spiking LIF behavior. Because of the remarkable correspondence between the electrical model of leaky integration of membrane voltage and the carrier density dynamics in an SOA, the temporal integration of the input optical pulses is completed by an SOA-based integrator. A subsequent NOLM with a Ge-doped fiber is employed as an ultrafast optical thresholder to unify the magnitude of pulses above the threshold, while the pulses below the threshold are eliminated. However, the integrator of this prototype could not respond to both excitatory and inhibitory inputs and could only output inverted integration results, which increase the system complexity. Moreover, the use of an optical pulse train as a spiking source compromises the asyn-
chronous response feature of a spiking neuron, and consequently limits the processing capacity at the repeating rate of the pulse train. Therefore some improvements were done on the elemental parts to solve theses issues.

To enable the temporal integrator to integrate both excitatory and inhibitory inputs, we studied the gain pumping effect in a C-band SOA and its simultaneous operation with gain depletion. The experimental results showed the feasibility of using a C-band SOA as a versatile integrator for both types of inputs, although a complete spiking neuron could not be demonstrated due to the lack of a proper laser source and modulator operating in the S-band. Another scheme by using XAM in an EAM was also proposed and demonstrated to achieve non-inverted integration, based on which an asynchronous spiking neuron was demonstrated. In the asynchronous response, instead of a pulse train, the spiking source is a duplicate of input pulses at another wavelength generated by FWM. In this way, the spiking source is only synchronized with the input pulses and the possibility of triggering multiple pulses by one input pulse is eliminated as well.

In addition, we demonstrated and studied power transfer function improvement and thresholding capability enhancement of the NOLM thresholder by using optical offset-spectral filtering. With the proposed Ge-doped NOLM with BPF configuration, the power transfer function is significantly improved, in terms of the steepness of the slope, as well as the flatness of the high and low levels.

Finally, two neuromorphic circuits based on our enhanced photonic neurons were presented. First, a signal feature recognition device inspired by the escape response of a crayfish was demonstrated. By two EAM-based integrators and one NOLM thresholder, the device was configured to accomplish fast and accurate feature recognition of three-bit input patterns at the picosecond level. Second, based on the gain pumping and gain depletion in a single SOA integrator and a NOLM thresholder, we demonstrated a timing jitter insensitive optical logic gate, which can be
reconfigured between AND and NOT operation. The slow recovery process of gain pumping/depletion is utilized to make the logic operation insensitive to timing jitter. By changing the threshold level, the timing-jitter tolerance window can be tuned. The measured timing jitter tolerance was up to $\pm 50$ ps for AND operation and $\pm 25$ ps for NOT operation.
Chapter 4

Supervised Learning in Photonic Neurons

Neurons communicate through electro-chemical spikes that are routed through a network of axons to cell bodies where they are integrated to trigger the firing of more spikes. This method of signal processing is both computationally efficient and scalable, adopting the best features of both analog and digital computing. Mimicking spike processing with photonics can result in bandwidths that are billions of times higher than biological neurons and substantially higher than electronics. One of the most powerful capabilities of neurons is their ability to learn, which takes place by strengthening synaptic connections in response to spiking activity. Hebbian learning rules \cite{70} states that connections between neurons are established based on correlations between pre-synaptic and post-synaptic activity. In other words, “Neurons that fire together wire together.” Modifying synaptic strength based on correlation with firing is called “Spike Timing Dependent Plasticity” \cite{14,71,72} or “STDP”. In this chapter, we report for the first time the demonstration of supervised learning based on STDP using photonic technology \cite{15}, potentially laying the foundation for learning at speeds a billion times faster than in biological neurons.
4.1 Hebbian learning rules and spike timing dependent plasticity

Perhaps the greatest strength of neurons is their ability to learn. Learning based on experience is essential for a system to dynamically adapt to an unpredictable or changing environment, or one that is too complex to be characterized a priori. There are several different methods by which neurons can learn: unsupervised learning extracts structure from an unknown signal environment for efficient encoding, dimensionality reduction, or clustering; supervised learning programs a system to perform a function defined in terms of its input-output pairs; and reinforcement learning optimizes a particular function based on a performance criterion. In all of these cases, the learning mechanism configures a set of system parameters, such as synaptic weights, based upon the activity of the system in response to its inputs, and as a consequence, the system function changes. “Synaptic plasticity” is not only responsible for learning and information storage but also the formation and refinement of neuronal circuits during development. With photonic neurons, we have focused on learning through synaptic plasticity, whereby the strength of input connections is determined by the relative timing of pre-synaptic and post-synaptic spikes.

Hebbian learning rules constitute a type of unsupervised learning in which changes to the strength of connections between two neurons (artificial or biological) is based on correlations between their pre-synaptic and post-synaptic activity. The rule, often paraphrased, “Neurons that fire together wire together,” adjusts weights so that the connections between nodes better represents the relationship between the activities of those nodes. The standard model for learning in systems of spiking neurons is a temporally-asymmetric form of Hebbian learning called, “Spike timing dependent plasticity” (STDP). STDP is a learning mechanism where the generated neuronal responses are based on the input timing, order, and sequence. For example, STDP rules
Figure 4.1: Schematic illustration of learning with two neurons, neuron 1 and neuron 2, with spike time dependent plasticity mechanism. $W_{1-2}$: Synaptic weight representing the synaptic strength of the connection between neuron 1 and neuron 2; $T_{1-2}$: Time delay representing the time for a spike from neuron 1 to arrive neuron 2; STDP: Spike-time dependent plasticity mechanism. (a) Unsupervised learning with STDP, the change in weight depends on the time interval between the post-synaptic spike and the pre-synaptic spike. (b) Supervised learning with STDP, the change in weight depends on the time interval between the post-synaptic spike and the teacher. (c) STDP characteristic. Horizontal axis represents the time interval $t_{\text{post}} - t_{\text{pre}}$ between the post-synaptic and pre-synaptic spikes. Vertical axis represents the resultant change in synaptic strength ($\Delta W$). Dark blue pulse: post-synaptic spike. Light pink pulse: pre-synaptic spike. Left hand side of the plot corresponds to post-pre firing, where pre-synaptic spike follows post-synaptic spike, the STDP mechanism causes a depression of the synaptic connection. Right hand side of the plot (shaded) corresponds to pre-post firing, where pre-synaptic spike precedes post-synaptic spike, the STDP mechanism causes a potentiation of the synaptic connection.

have been applied to coincidence detection \cite{76}, sequence learning \cite{77} \cite{78} \cite{79}, path learning in navigation \cite{80} \cite{81}, and directional selectivity in visual response \cite{81} \cite{82}. Figure 4.1(a) and (b) show the architecture of a neuron using STDP for unsupervised and supervised learning, respectively. In STDP, the relative timing of pre- and post-synaptic spikes determines the changes to the strength of the synapse as follows:
- Pre-post firing: a pre-synaptic spike precedes a post-synaptic spike and is presumed to have contributed to causing the post-synaptic spike. Therefore, the synaptic connection is “rewarded” or strengthened, causing a “potentiation”.

- Post-pre firing: a pre-synaptic spike follows a post-synaptic spike and could not have contributed to causing the post-synaptic spike. Therefore, the synaptic connection is “punished” or weakened, causes a “depression”.

Figure 4.1(c) illustrates STDP behavior, where the shaded part is the potentiation window while the un-shaded part is the depression window. Previous research has shown that the width of the potentiation and depression windows varies asymmetrically at different synapses [83].

### 4.2 Optical spike timing dependent plasticity

To enable learning by photonic neurons, this section presents the first demonstration of STDP using optical components [15] by Dr. Mable P. Fok. The operation of optical STDP is similar to its physiological counterpart, in which the strength of the optical interconnection is increased for pre-post firing while it is decreased for post-pre firing. The effect is stronger when the pre- and post-synaptic spikes are closer, and weaker otherwise. Unlike biological STDP, optical STDP operates on picosecond time scales, enabling learning at speeds nine orders of magnitude faster than in biological neurons. This is achieved by exploiting fast physical mechanisms in photonic devices, such as absorption saturation in electro-absorption modulators (EAM) [84] and gain depletion in semiconductor optical amplifiers (SOA) [85].

The optical implementation of STDP in the photonic neuron is illustrated in Figure 4.2(a). Post- (dark blue) and pre-synaptic (light red) spikes are combined and split unequally into two branches. The upper branch consists of strong post-synaptic
Figure 4.2: Operation principle of optical spike time dependent plasticity (STDP)
(a) Optical implementation of the physiological STDP mechanism. (b) Formation of depression window in SOA. \( t_{\text{pre}} \): time where pre-synaptic spike is launched; \( t_{\text{post}} \): time where post-synaptic spike is launched. (c) Formation of potentiation window in EAM. (d) Linearly combining the SOA and EAM effect results in STDP characteristic. (e) Experimentally measured optical STDP characteristic. \( \Delta W \): change in weight, \( t_{\text{post}} - t_{\text{pre}} \): time interval between the post- and pre-synaptic spikes.
spike and weak pre-synaptic spike is launched to the SOA (Kamelian SOA-NL-L1-C-FA) that is driven by a current source; the lower branch consists of strong pre-synaptic spike and weak post-synaptic spike is launched to the EAM (CIP 10G-PS-EAM-1550) that is negatively biased by a voltage source. The post-synaptic spike induces gain depletion in the SOA and the resultant gain is experienced by the pre-synaptic spike. The pre-synaptic spike induces absorption saturation in the EAM and the resultant absorption is experienced by the post-synaptic spike. Optical bandpass filters (Filter 1 and Filter 2) are used to extract the pre- and post-synaptic spikes from the SOA and EAM outputs, respectively. They are then detected using a photodetector. Figure 4.2(b) and (c) show the mechanism for optical STDP. When a strong post-synaptic spike enters the SOA, it immediately depletes the SOA gain. The gain gradually recovers over time due to the presence of the SOA driving current. Therefore, if a pre-synaptic spike enters the SOA shortly after the post-synaptic spike [Figure 4.2(b)(i)], it experiences the minimum gain, i.e., the weakest output power. As the gain recovers over time, the output power of the pre-synaptic spike after the SOA gradually increases as the time interval between the post- and pre-synaptic spikes increases, forming the depression window [region I in Figure 4.2(b)(iii)]. If a pre-synaptic spike enters the SOA before the post-synaptic spike [Figure 4.2(b)(ii)], it does not experience any gain depletion from the post-synaptic spike, and the output power will remain the same despite the time interval between the post- and pre-synaptic spikes [region II in Figure 4.2(b)(iii)]. In the EAM, the absorption is large due to the reverse voltage bias, and no absorption saturation is experienced by the post-synaptic spike when it enters the EAM before the strong pre-synaptic spike [Figure 4.2(c)(i)]. Thus, a weak output power is obtained despite the time interval between the pre- and post synaptic spike [region III in Figure 4.2(c)(iii)]. When the post-synaptic spike enters the EAM shortly after the pre-synaptic spike [Figure 4.2(c)(ii)], a strong absorption saturation effect (weak EAM absorption) is experienced by the post-synaptic
spike, resulting in a strong output power. The EAM absorption builds up gradually over time due to the reverse voltage bias, resulting in a gradual decrease in output power of the post-synaptic spike, forming the potentiation window [region IV in Figure 4.2(c)(iii)]. By linearly combining the effects in the SOA and EAM, an STDP characteristic is obtained, as plotted in Figure 4.2(d).

Figure 4.3: Reconfigurable STDP characteristic, ∆W: change in weight, $t_{post} - t_{pre}$: time interval between the post- and pre-synaptic spikes. (a) Stronger EAM reversed bias voltage results in decrease in potentiation window width; (b) Increase in SOA driving current results in decrease in depression window width; (c) Decrease in pre- and post-synaptic spikes splitting ratio results in increase in potentiation window height and decrease in depression window height; (d) Decrease in input power to the EAM decrease the potentiation window height.

In the experiment, the input powers of the pre-synaptic and post-synaptic spikes are both 10 dBm with ~5-ps pulse widths. Figure 4.2(e) shows the experimentally measured STDP responses using the proposed optical STDP configuration, which
resembles the physiological STDP characteristic. As with biological STDP, optical STDP allows independent control of the width and height of the potentiation and depression windows to implement different functions. The potentiation window width decreases as the EAM bias increases [Figure 4.3(a)], resulting from the stronger electro-absorption effect under a stronger bias, while the depression window width decreases as the SOA current increases [Figure 4.3(b)], due to the increase in carrier density in the SOA under a stronger injection current. The ratio between the potentiation and depression window heights changes as the pre- and post-synaptic spikes splitting ratio varies [Figure 4.3(c)]. In the SOA, a larger splitting ratio means that a stronger post-synaptic spike is used to induce stronger gain depletion, resulting in further decreases in power of the pre-synaptic spike, i.e. height of the depression window increases. In the EAM, however, although a strong pre-synaptic spike is used to induced stronger absorption saturation, the maximum output power of the output post-synaptic spike as well as the potentiation window height is limited by the post-synaptic spike power launched to the EAM, which decreases as the splitting ratio is larger. The height of the potentiation window is adjustable by independently changing both the pre- and post-synaptic spikes power launch to the EAM [Figure 4.3(d)]. A decrease in the power leads to a decrease in window height that results from the power limit set by the post-synaptic spike as well as weaker absorption saturation. By individually adjusting the pre- and post-synaptic spikes power, both the height and width of the potentiation and depression windows are variable. Supplementary Table 4.1 summarizes the effect on STDP windows resulting from an increase in various parameters in optical STDP.
Table 4.1: Effects on STDP window width and height when the following parameters are increased

<table>
<thead>
<tr>
<th>Potentiation window</th>
<th>EAM bias</th>
<th>SOA current</th>
<th>Splitting ratio</th>
<th>Pre-synaptic spike power</th>
<th>Post-synaptic spike power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Width</td>
<td>⇓⇓</td>
<td>-</td>
<td>-</td>
<td>⇑</td>
<td>-</td>
</tr>
<tr>
<td>Height</td>
<td>⇑</td>
<td>⇧</td>
<td>⇦</td>
<td>⇧</td>
<td>⇧</td>
</tr>
</tbody>
</table>

| Depression window   | Width    | ⇧           | ⇦               | ⇧                        | ⇧                        |
| Height              | ⇦         |             | ⇧               | ⇧                        | ⇧                        |

4.3 Supervised learning in a photonic neuron based on optical STDP

Using the proposed optical STDP in the previous section we experimentally demonstrate supervised learning of a photonic neuron in which a “teacher” determines the way the photonic neuron should spike in response to its inputs [15]. This powerful feature of photonic neurons opens up a wide range of applications such as adaptive control, perception, sensory processing and autonomous robots.

Figure 4.4 illustrates the schematic of supervised learning in a photonic neuron based on optical STDP feedback control. In supervised learning, the output from the neuron is compared with a “teacher’s” input instead of the neuron’s input. In the learning phase, the input training pulse patterns are sent through one input branch, where they are applied a certain weight value and time delay and then reach the photonic neuron. After the spiking processing in the neuron, including temporal integration and optical thresholding, the output from the neuron is fed into the post-synaptic input of the STDP module. Simultaneously at the pre-synaptic input port, a “teacher” inputs a string of teaching patterns, which are the correct response that the neuron should spike out with the input training patterns. Then the STDP module compares the neuron’s output with the teacher input patterns. If the neuron is not
responding in the way the “teacher” wants, the STDP module, according to the characteristic curve in the subset of Figure 4.4, will send out a change signal to the weighting unit at the neuron’s input to alter the interconnection strength (weight value). This compare-and-change process keeps going until the neuron is spiking exactly the same as the teacher input; the STDP will then no longer change the weight value and the learning phase is over. After the learning phase, the “teacher” is removed or the feedback loop is cut off from the weighting unit, and consequently the neuron fixes the interconnection strength and is ready for use.

Figure 4.4: Supervised learning in a photonic neuron based on adaptive feedback control loop. Subset: the setup and transfer curve of the optical STDP module.

Figure 4.5: Experimental setup of supervised learning with a photonic neuron and optical STDP. EAM: electro-absorption modulator; EOM: electro-optic modulator; NOLM: nonlinear optical loop mirror; PD: photodetector.
The experimental setup is illustrated in Figure 4.5. In the experiment, the training input at 1551.72 nm consists of three patterns ‘1011’, ‘1101’ and ‘0110’, labeled as pattern \(A\), \(B\) and \(C\). Every ‘1’ is a 9-ps wide optical pulse, and the adjacent bits in the same pattern are 30 ps away from each other. The time interval between two adjacent patterns is 2 \(\mu\)s and the three patterns are repeated periodically in the sequence of ‘\(ABC\)’. The weighting unit is an electro-optic modulator (EOM) attenuating input patterns according to the applied DC bias voltage. The photonic neuron setup we used here is an asynchronous neuron as in reference [48], and a sampling pulse train at 1553.33 nm is generated by duplicating the training pulses input to the neuron. The temporal integration in the neuron is realized by cross-absorption modulation of an inverted-biased EAM, which is biased at \(-1\) V to enable a \(\sim 150\)-ps integration window. After the integrator, an optical filter centered at 1553.33 nm passes only sampling pulses as the integration results. The results are amplified by an Erbium-doped fiber amplifier (EDFA) for optical thresholding in the next step, before which out-of-band amplified spontaneous emission (ASE) noise gets suppressed by another optical filter. The optical thresholding is based on a nonlinear optical loop mirror with a 12-m long Ge-doped fiber [40].

After optical thresholding, the output from the neuron and teacher input are equalized to the same power and synchronized in time by EDFAs and time delay lines respectively. In the experiment, teacher input patterns are generated as ‘0011’, ‘0101’ and ‘0010’, labeled as patterns \(A'\), \(B'\) and \(C'\), corresponding to input patterns \(A\), \(B\) and \(C\). In the optical STDP module, patterns \(A\), \(B\), \(C\) are compared with \(A'\), \(B'\), \(C'\) respectively. Since the output from the optical STDP module are optical pulses, which do not have negative power values, there is actually an offset lifting up the whole transfer curve in the subset in Figure 4.4 above zero. Therefore an operational amplifier (Op-Amp) circuit in the O/E conversion module subtracts the offset value from STDP’s output value and holds the DC voltage that is fed back to
control the weighting unit. It is worth noting that in the optical STDP based on an SOA and EAM, introduced in Section 4.2, the offset value always comes from the pre-synaptic pulses amplified by the SOA, as shown in the subset in Figure 4.4. Thus in this case the pre-synaptic input (teacher input) is subtracted in the O/E conversion module. In the experiment, the pulse power of the neuron’s output and teacher pulses is equalized at $\sim 10$ dBm before launching into the STDP module. The EAM and SOA in the STDP module are biased at $-4$ V and 70 mA to get the transfer function shown in Figure 4.4.

Figure 4.6: Operation principle of the learning phase in supervised learning.

The operational principle of the learning process is illustrated in Figure 4.6. In the first row are the input patterns $A$, $B$ and $C$ to the neuron. Inside the neuron, the leaky-integration dynamic is illustrated by the profile curves in the second row. Thus at each input pulse the integration value is sharply raised up as a peak. The pink sampling pulses are attenuated according to the dynamic profile as shown in the second row as well. The red dashed line shows the threshold level, and every sampling pulse above this level will pass through the following optical thresherolder as output spikes. Focusing first on the left column in Figure 4.6 when the weight value is correct, the first pulse in pattern $A$, $B$ or $C$ is not large enough to make the neuron spike. But after the second pulse in each pattern contributes, the integrated results are above the threshold level to trigger output spikes. Therefore the neuron’s output,
the green spikes shown in the third row in Figure 4.6 is two spikes for pattern A, two for pattern B, and one for pattern C. In the forth row are the teacher pulses as pre-synaptic input for the STDP module. Since the weight value is correct, the output patterns are exactly same as the teacher input patterns $A'$, $B'$, and $C'$. In the experimental setup, the teacher input is precisely aligned with the neuron’s output in time. Taking the first pattern of the neuron’s output and teacher input as an example, we mark the two output pulses from the neuron as ‘a’ and ‘b’, and two teacher pulses as ‘$a'$’ and ‘$b'$’. According to the characteristic curve of STDP, $a$ and $a'$ are aligned in time, so they do not generate any output via STDP. The same holds true for $b$ and $b'$. Moreover, since $a$ is leading $b'$ and $b$ is lagging $a'$, the STDP module generates ‘decrease’ and ‘increase’ feedback signals, respectively. However, noticing that the lead and lag times are both 30 ps, according to the transfer curve in Figure 4.4, the output amplitudes are the same as well, but with opposite signs. Therefore, at the output of STDP, the positive and negative outputs cancel, and do not send out any change signal to the weighting unit, as shown in the fifth row in Figure 4.6.

When the weight value is initially too low, the results are illustrated in the middle column in Figure 4.6. For input pattern A, it is attenuated so much that the first and second pulses are not strong enough to trigger spiking, but only after integrating all three input pulses does the integration result exceed the threshold level and releases an output pulse. The neuron output thus has only one pulse ($b$) instead of two ($ab$) for pattern A in this case. Since the first two pulses in pattern B (or C) are closer than in A, their integration is still large enough to make the neuron spike. Thus the output for B (or C) is still two (or one) pulses, same as in the previous case with correct weight value. Consequently, the first teacher pattern $A'$ is different from the neuron’s output, since $a$ is missing. Then in STDP, pulse $a'$ and $b$ constitute a pre-post firing and feedback an ‘increase’ signal to the weighting unit.
In the case that the weight value is too high, as shown in the right column in Figure 4.6, every single input pulse is so strong that the neuron spikes for each of them. Therefore, at the neuron’s output port, every pattern has an additional output pulse compared with teacher patterns. As illustrated in the third and forth rows, the additional pulses with teacher pulses create post-pre firings, which generate ‘decrease’ signal feedback to weighting unit. In conclusion, no matter whether the initial weight is too low or too high, the weight value will be brought back to the expected value after several iterations through the STDP and feedback loop. Therefore the learning process is achieved. After the learning phase, by simply tuning off the teacher pulses, the weight value is fixed and the neuron’s behavior maintained.

Figure 4.7: Experimental results of the learning phase of supervised learning.

Figure 4.7 illustrates the experimental results. Shown on the top are the input patterns A, B and C, which are attenuated by an EOM as the tunable weight unit, providing a 30-dB tuning rage for input pulse power. In the experiment, the average power of the input patterns is tuned from $-7$ dBm (low weight) to 15 dBm (high weight). For different weight values (different columns), the integration results output from the integrator are shown in the second row in Figure 4.7. Since the transfer function for an optical thresholder based on NOLF is fixed, the threshold level can be tuned by adjusting the input pulse power, i.e. the gain of the EDFA. In this
experiment, the gain is set to about 20 dB to get a threshold level shown by the dashed line illustrated in Figure 4.7. After thresholding, the neuron’s outputs are shown in the third row, with all weak pulses below threshold removed and strong pulses above threshold unified in amplitude. For different weight values, the experimental results are consistent with the theoretical analysis in Figure 4.6. At the bottom of Figure 4.7 are the teacher patterns $A'$, $B'$, and $C''$, corresponding to the neuron’s output with correct weight values.

During the learning phase, when the initial weight value is set too low or too high, the optical STDP module starts to generate feedback signals. The output pulses from the optical STDP and teacher’s input pulses are detected by the PDs and then sent into the Op-Amp integrator circuit in the O/E conversion module. The Op-Amp circuit, which includes a TL081CP chip in a simple differential integrator setup, subtracts the offset (power of teacher’s pulses) from STDP’s output, integrates the pulse signals of positive or negative weight changes and holds the DC output as the control voltage for the weighting unit. The training sequence ‘$ABC$’ and teacher sequence ‘$A'B'C''$’ are repeatedly sent into the photonic neuron and the STDP module respectively. As shown in Figure 4.7 after iterations of feedback from the STDP, the weight value converges to the correct value. After the learning phase is done, the feedback loop or the teacher input can be cut off to fix the weight value, and consequently the photonic neuron will work as if it has learned from the teacher.

The convergence speed of the learning process is determined by the repetition frequency of the teacher patterns and the increment/decrement amount (step size) of the weight value in each iteration. In the experiment, the TL081CP chip used in the Op-Amp circuit has a bandwidth of 3 MHz, which limits the repetition frequency of teacher patterns below 3 MHz. By using a higher-speed Op-Amp chip and fast enough PDs, the teacher patterns can be repeated with up to GHz frequency. Meanwhile, the upper bound of the repetition rate will be imposed by the integration window (150
ps in this experiment) of the photonic neuron, since the integration value has to be reset to zero between adjacent patterns. Regarding the step size in each iteration, it is equal to the product of output power from the STDP module and the gain of the Op-Amp circuit. Large step size may cause over-tuning and oscillation around the convergence point, while a small step size may increase the sensitivity to noise and slow down the convergence speed. In our case, the convergence speed is also affected by the relatively large latency due to a long feedback path. The whole feedback path includes several spans of patch cords and time delay lines for pulse synchronization and three EDFAs, which introduces ~600-ns latency to the weight unit. However, since the photonic neuron and optical STDP work at picosecond speed, the large latency does not affect the capacity or the speed of the photonic neuron itself. Also, by replacing EDFAs with SOAs and integrating the components onto a single chip, the latency can be greatly decreased to a few nanoseconds.

Figure 4.8: Photos of the bench-top supervised learning setup: (a) Tunable weighting unit for input pulse train, and pattern construction of teacher, sampling and input pulse patterns; (b) Temporal integrator, thresholder and STDP modules.

Figure 4.8 shows the photos of the bench-top supervised learning setup. In Figure 4.8(a) the EOM acting as a tunable weighting unit is placed in the middle. On the left are power splitters, the attenuators and fiber delay lines to construct a de-
sired pulse pattern as teacher pulses. On the right is a similar setup for sampling
and input pulses. Since the sampling and input pulses use the same pattern, they
are split from the same pulse pattern and then weighted and delayed separately. In
Figure 4.8(b), on the left is the optical STDP module introduced in the last section,
and at the top right and bottom right corner are the EAM integrator and Ge-doped
NOLM thresholder, respectively.

4.4 Conclusion

In summary, we proposed and successfully demonstrated, for the first time, optical
STDP and supervised learning for photonic neurons. This is an important mile-
stone enabling photonic neurons to perform learning. This photonic neuron is based
on a well-established LIF model where neurons communicate using spikes. Based
on spike processing, low-noise and high-efficiency signal processing can be achieved.
STDP modifies synaptic strengths based on the correlation between the pre- and
post-synaptic activities, allowing the neuron to learn according to the environment or
a teacher. The photonic neuron learned to spike in the way expected by the teacher
through STDP-based supervised learning. Our optical STDP utilized the inherent
properties of two optical devices, an electro-absorption modulator and a semiconduc-
tor optical amplifier, that naturally mimic a biological STDP characteristic. Just
like in its biological counterpart, variable potentiation and depression windows are
achieved in optical STDP, satisfying the requirements of different types of synapses
and applications. The above demonstration forms the foundation for a photonic neu-
ron to learn about the environment like a biological brain, while having a processing
speed that is a billion times faster.
Chapter 5

Phase-Sensitive Boosting for Long-Haul WDM Systems using Optical Phase Conjugation

Besides unconventional and high-speed computing, another area requiring ultrafast signal processing technology is telecommunications. Nowadays, sophisticated coherent detection and digital signal processing (DSP) technologies have pushed the single channel transmission capacity up to more than 1 Tb/s in long-haul transmission systems [16]. It has thus been extremely difficult for electronic devices to accomplish real-time processing at such a high operational speed. Therefore researchers have been turning to optics to address this issue as well. In this chapter, we propose and demonstrate a scheme using all-optical phase conjugation by four-wave mixing (FWM) to mitigate the nonlinear impairments in coherent ultra-long-haul transmission systems [17] [18].
5.1 Introduction

Coherent detection and DSP have been the enabling technologies for improving spectral efficiency in long-haul wavelength-division multiplexing (WDM) transmission using polarization multiplexed multi-dimensional modulation formats \[86, 87\]. Since linear impairments such as chromatic dispersion (CD) and polarization mode dispersion (PMD) can be compensated by DSP \[88\], accumulation of amplified spontaneous emission (ASE) noise from inline erbium-doped fiber amplifiers (EDFAs) and fiber nonlinear impairments are the limiting factors to system reach. Fiber nonlinear impairments can be categorized into single-channel effects created by self-phase modulation (SPM), and effects due to WDM operation, such as four-wave mixing (FWM) and cross-phase modulation (XPM). It has been reported that fiber dispersion in dispersion-unmanaged links can substantially reduce SPM-induced nonlinear penalty \[89, 90\]. In dispersion-managed transmission, however, SPM still imposes a severe limitation on signal launch power and therefore the transmission distance \[89, 90, 91\]. To increase the nonlinearity tolerance of coherent systems, several optical and electronic approaches have been reported. Electronic pre-compensation \[92, 93, 94\], post-compensation \[95, 96, 97, 98\] and the combination of both \[99, 100\] have been used to pre-distort and correct the signal phase according to the signal power variation, which are subject to the accuracy of the link nonlinearity estimation. Digital back-propagation is attractive for its flexibility and performance \[101\]. However it requires \textit{a priori} knowledge of the systems dispersion map and optical power map, as well as extensive computational resources. Mid-span phase conjugation employs optical \[102\] or electrical devices \[103\] to conjugate the optical signal phase at the middle point of the transmission link in order to achieve cancellation of the nonlinear phase shifts from the two link halves. To obtain meaningful performance improvement using the scheme, however, not only does the whole link need to be homogeneous, but also the
signal power evolution profiles before and after mid-span phase conjugation need to resemble as mirrored images.

Optical 2R regeneration using frequency-degenerate and non-degenerate phase sensitive amplifiers (PSA) \([104][105][106][107][108][109]\) has been proposed to increase optical signal-to-noise ratio (OSNR) and mitigate the nonlinear distortion. In degenerate PSAs \([104][105][106]\), two optical pumps symmetrical (or identical) to the phase-modulated signal in frequency are used to generate an idler with identical frequency as signal via FWM, and through coherent sum of the signal and idler, the signal phase is stabilized with respect to the pumps. Non-degenerate PSAs \([107][108][109]\), on the other hand, employ FWM of single pump with signal and idler at different frequencies to stabilize the sum of signal and idler phases. However, degenerate PSAs are sensitive to the modulation format and can only be applied on single wavelength for a fixed pump configuration, which limits their applicability in multi-phase modulation formats and WDM systems. Moreover, due to stringent requirements in aligning the phase of the pump and signal at the PSA, both degenerate and non-degenerate PSAs require high-quality carrier regeneration and optical phase locking, which dramatically increases system complexity. It is worth noting that to maintain a system noise figure (NF) lower than the traditional long-haul systems and improve the performance to the ideal value of 6-dB gain \([107]\), all in-line amplifiers would have to be replaced by PSAs, otherwise the overall NF will be dominated by the phase-insensitive amplifiers in the system. Unfortunately replacing in-line EDFAs with PSAs is impractical because of system complexity.

In the following sections of this chapter, we propose and demonstrate a hybrid optical/digital scheme based on optical phase conjugation and DSP to improve OSNR and nonlinear tolerance in long-haul transmission systems, by both simulation and experiment. FWM is adopted to generate an idler, which is a phase-conjugate copy of the original optical signal. The signal and phase-conjugated idler are transmitted
together and detected separately using two coherent receivers. In the DSP module, the idler symbols are conjugated digitally and summed with the signal symbols with proper phase alignment, thereby suppressing the nonlinear phase shifts contributed from both signal and idler components by phase-sensitive boosting (PSB). Moreover, by combining the signal and idler components, the SNR per symbol can be enhanced digitally by 3 dB at linear regime in principle. The proposed hybrid optical/digital scheme is transparent to modulation format and optical fiber types, and does not require optical phase locking or carrier regeneration thanks to digital frequency offset compensation in DSP. Obviously the phase-conjugate copy consumes additional bandwidth over the original optical signal, so the PSB scheme trades spectral efficiency for longer system reach. Therefore it is much more suitable for systems which demand ultra-long system reach and high signal quality but with less requirement on spectral efficiency. It can also be applied at a network node where bandwidth would be available for implementing PSB due to the reduced number of output channels/wavelengths after the traffic divergence (i.e. branching unit in a submarine network).

5.2 Principle of digital phase-sensitive boosting technology

Figure 5.1 shows the setup of our proposed optical/digital PSB scheme applied to a dual-polarization quadrature phase-shift-keying (DP-QPSK) system [17]. To generate the phase-conjugated copy, we combine the DP-QPSK signal with a high-power continuous-wave (CW) pump whose polarization is positioned at 45° relative to both orthogonal polarizations of the polarization-multiplexed signal. A polarization beam splitter (PBS) aligned with the signal polarization axes de-multiplexes the polarization components and projects the pump equally onto the two axes. In each branch, a
highly nonlinear fiber (HNLF) is used to generate a phase-conjugated idler via FWM. Assuming the signal and pump phases are $\theta_s$ and $\theta_p$, respectively, the idler phase $\theta_i$ satisfies the phase relation:

$$\theta_s + \theta_i = 2\theta_p$$

(5.1)

Given that we can define $\theta_p = 0$ for the pump phase, the idler phase, $\theta_i = -\theta_s$, is a phase-conjugate copy of the signal. A second PBS recombines the two polarization components of the signal and idler. An optical filter then rejects all unwanted frequencies outside of the signal and idler bands, which are then amplified and transmitted together.

Figure 5.1: Operation principle of digital phase-sensitive boosting. EDFA: erbium-doped fiber amplifier, HNLF: highly nonlinear fiber, PBS: polarization beam splitter.

During transmission, both the signal and idler experience SPM and XPM. Provided the signal and idler are relatively close in wavelength so they experience near-identical CD, the nonlinear phase-shift $\theta_{NL}$ exerted on both the signal and the idler are near-identical as well. At the receiver, the signal and idler phases can be expressed as $\theta_s + \theta_{NL}$ and $\theta_i + \theta_{NL}$, respectively. Two coherent receivers are used to downconvert the signal and idler to electrical baseband. CD, PMD, frequency offset and laser phase noise are compensated in DSP. Assuming that phase error due to carrier recovery is negligible, we can conjugate the phase of the idler symbols digitally and obtain $-\theta_i - \theta_{NL}$. The signal and conjugated idler component are then digitally summed up with proper timing and phase alignment to obtain:
\[ e^{\theta_1 + \theta_{\text{NL}}} + e^{-\theta_1 - \theta_{\text{NL}}} = e^{\theta_1 + \theta_{\text{NL}}} + e^{\theta_1 - \theta_{\text{NL}}} = 2 \cos \theta_{\text{NL}} \cdot e^{\theta_1} \] (5.2)

It is observed that the nonlinear phase-shift results only in amplitude fluctuation of magnitude \(2 \cos \theta_{\text{NL}}\). Due to the positions of the nearest neighboring points in the QPSK constellation, the information in \(\theta_s\) is much less affected by the amplitude fluctuation in the radial direction, as shown by the inset in Figure 5.1. Note that, by digitally compensating the frequency offset and phase noise between signals and local oscillator (LO) lasers, the use of complicated all-optical carrier recovery techniques used in most PSA demonstrations [104][105][106][107][108][109], is no longer required.

In addition to single channel application, the proposed PSB scheme can be extended to WDM systems. Similarly to the setup in Figure 5.1, multiple wavelength channels can share the same optical pump and locate on one side of the pump. Then after FWM, the idlers will be generated on the other side of the optical pump, in reverse spectral order compared with the original signals. To keep the validity of near-identical CD between each pair of signal and idler, dispersion-managed links are preferred for WDM applications.

### 5.3 Numerical simulation

As a proof of concept, we conducted a numerical simulation of the proposed digital PSB scheme using VPItransmissionMaker™, based on the setup in Figure 5.1 [18]. A 112-Gb/s DP-QPSK signal, modulated by pseudo-random binary sequence (PRBS) length of \(2^{15}-1\) at four parallel lanes, is launched into the optical conjugate copier. The signal wavelength and pump wavelength are 1550.1 nm and 1549.3 nm respectively, at 100-GHz spacing. Pump power is set to 1 W to generate an idler at 1548.5 nm with the same power level as the original signal. To study the performance of the proposed PSB scheme on ultra-low-loss fiber transmission for future
link installations, ultra-low-loss fiber links and parameters are used in the simulation. The transmission link consists of multiple spans of 60-km ultra-low loss fiber with EDFAs at 5-dB NF to compensate span loss. The fiber attenuation and nonlinear coefficient (\( \gamma \)) are 0.161 dB/km and 0.731 W\(^{-1}\)/km. The channel launch power is maintained at −1 dBm for both signal and idler. After transmission, the signal and idler components are separately detected by two coherent receivers with LO set at 1550.1 nm and 1548.5 nm. Then the digitized signal and idler are processed with DSP, where we use a frequency-domain equalizer (FDE) to compensate CD, and an adaptive time-domain equalizer (TDE) to compensate other linear impairments. After frequency offset compensation and carrier recovery, the equalized symbols in the idler channel are phase-conjugated and recombined with the signal symbols. The resulting symbols are then analyzed to obtain the signal Q-factor. We tested the proposed PSB scheme in both dispersion-managed and dispersion-unmanaged links. In the dispersion-managed link, dispersion compensation fiber (DCF) is deployed at the end of each span to completely compensate the span dispersion, while in the dispersion-unmanaged link the accumulated dispersion in the whole link is compensated by DSP at the receiver end. The split-step Fourier method [110] is implemented in VPItransmissionMaker\textsuperscript{TM} for the transmission simulation by solving the nonlinear Schrödinger equation. The simulation results are plotted in Figure 5.2.

As shown in Figure 5.2(a) and (b), for both dispersion-managed and dispersion-unmanaged links with −1-dB channel power for both signal and idler, the proposed digital PSB-based transmission scheme can improve the Q-factor by 2.5 dB after 6300-km transmission, due to its enhanced noise and nonlinearity tolerance. At shorter transmission distances, the nonlinear effect is relatively weak, so the dominant distortions are due to the noises from transmitters and receivers, which cannot be corrected by PSB. Therefore the Q-improvement over the original signal increases with the transmission distance because of less received OSNR and more nonlinear distortions.
Figure 5.2: Q-factor improvement by digital PSB in (a) dispersion-managed link and (b) dispersion-unmanaged link; Q-factor versus channel power comparison of PSB scheme, DSC scheme and original signal in (c) dispersion-managed link and (d) dispersion-unmanaged link.

To verify that the digital PSB scheme does suppress nonlinear distortion on top of improving sensitivity by doubling signal amplitude, we compare the performance of PSB with a direct signal-copying (DSC) scheme. To create the DSC signals, we remove the phase-conjugate copier in Figure 5.1 and then couple and modulate two laser outputs at 1550.1 nm and 1548.5 nm together using the same DP-QPSK transmitter. As a result, both signals will carry the same data for transmission through the link. At the receiver side, after digital equalization in DSP, the two signals are directly summed up without phase conjugation. Therefore, the DSC scheme takes advantage of doubled signal amplitude just like the PSB scheme, but without the suppression of phase noise distortion. By monitoring the Q-factor while sweeping the channel power, the nonlinear power tolerance has been enhanced by about 1-dB in
PSB over DSC in both dispersion-managed and dispersion-unmanaged links, as illustrated in Figure 5.2 (c) and (d). The maximum Q-factor is enhanced by 0.7 dB in dispersion-managed link [Figure 5.2(c)] and by 0.5 dB in dispersion-unmanaged link [Figure 5.2(d)]. The performance difference between two types of links is likely due to the fact that dispersion-unmanaged links are intrinsically more tolerant to nonlinear distortion than dispersion-managed links [89][90].

5.4 Experimental demonstration

In this section, the proposed scheme is experimentally demonstrated using a fiber recirculating loop comprising 860-km spans of dispersion-managed fiber (DMF), as shown in Figure 5.3 [18]. At the transmitter, an external cavity laser at 1550.12 nm (\(\lambda_2\)) is modulated with 28-Gbaud QPSK symbols using an I/Q modulator. The I and Q signals are generated from a 4:1 electrical multiplexing of a 7-Gb/s 2\(^{15} – 1\) PRBS pattern. The 56-Gb/s QPSK signal is coupled with a 26-dBm CW pump at 1550.52 nm and launched into a two-meter Bismuth oxide-based nonlinear fiber (Bi-NLF), which generates a phase-conjugated idler at 1550.92 nm (\(\lambda_4\)). The two-meter Bi-NLF has an attenuation of 5.2 dB, a nonlinear coefficient of 1050 W\(^{-1}/\)km, a dispersion of –250 ps/nm/km at 1550 nm and a dispersion slope less than 0.5 ps/nm\(^2)/km\) across C-band [111]. Thanks to the high stimulated-Brillouin-scattering (SBS) threshold (higher than 30 dBm) [112], an FWM efficiency of –11 dB is achieved without phase modulation of the CW pump for SBS suppression. An optical interleaver and a wavelength selective switch (WSS) are used to remove residual pump and to equalize the signal and idler powers. After equalization the OSNR of signal and idler is higher than 40 dB. The signal (\(\lambda_2\)) and idler (\(\lambda_4\)) form the even channels in this experiment, as shown in Figure 5.3. For the odd channels, we passively combine three lasers and modulate them with 28-Gbaud QPSK by the same technique using another separate
4:1 electrical multiplexer with decorrelated $2^{15} - 1$ PRBS inputs resulting in different patterns from the even channels. Polarization multiplexing is performed separately on even and odd channels by splitting their signals, delaying one copy by 224 symbols, and rotating it to the orthogonal polarization followed by polarization combining. The odd- and even-channels are then combined with a 50/100-GHz optical interleaver and launched into the fiber re-circulating loop together with eighty-five CW lasers, as illustrated by the spectrum in Figure 5.4(a). The loop has 860-km spans of dispersion-managed fiber (DMF) with an inline EDFA after each span. Each span is composed of a 40-km positive dispersion (19 ps/nm/km) fiber with 0.2 dB/km loss and effective area of 93 $\mu m^2$, and another 20-km negative dispersion ($-38$ ps/nm/km) fiber with 0.245 dB/km loss and effective area of 24.5 $\mu m^2$. The EDFA output power is fixed at 16 dBm.

At the receiver, the five modulated channels ($\lambda_1 - \lambda_5$) are extracted by a WSS and amplified by an EDFA, followed by a splitter and two synchronized coherent receivers. In each coherent receiver, the signal and idler are downconverted to electrical baseband by combining them with LO lasers centered at the appropriate frequencies.

Figure 5.3: Experimental setup of digital phase-sensitive boosting for 112-Gb/s DP-QPSK transmission. DMF: dispersion-managed fiber, IL: interleaver, OC: optical coupler, VOA: variable optical attenuator, WSS: wavelength selective switch.
using polarization-diversity 90° optical hybrids followed by balanced photodetectors. The electrical signals are then sampled by two synchronized quad-channel digitizing oscilloscope at 40-GSa/s and 16-GHz bandwidth. Then the same offline DSP as in the simulation is used to process the signal and idler. The bottom left inset in Figure 5.3 shows the constellations of the signal symbols before and after phase-sensitive boosting.

We firstly measured the back-to-back (B2B) performance without nonlinear impairment. Odd channels are turned off with only the signal and idler channels remaining. The signal Q-factor, derived from BER measurement, is monitored when amplified spontaneous emission (ASE) noise is added at the receiver. Figure 5.4(b) shows the Q vs. OSNR results. It is observed that phase-sensitive boosting improves the Q-factor by 2 to 2.4 dB, due to doubling of signal amplitude. At very high OSNR, the Q-factor improvement is somewhat less due to signal distortion by the transmitter which cannot be removed by digital PSB.

At 4,800 km (10 loops), we swept the launch power per channel (signal, idler and three odd-channels) from $-9$ dBm to $+1$ dBm. To show that the digital PSB scheme does not only improve sensitivity by doubling signal amplitude, but also suppresses nonlinearity, we compare the performance of PSB with the DSC scheme where we replace the idler at 1550.92 nm with a duplicate of the original signal (remove the HNLF and the pump laser in Figure 5.3 and insert a laser at $\lambda_4$ before the I/Q modulator). For DSC, the equalized symbols in $\lambda_2$ and $\lambda_4$ are added up without phase conjugation. The experimental results for PSB and DSC are shown in Figure 5.4(c). As can be seen, the PSB scheme is able to increase the optimum launch power from $-5.2$ dBm to $-4.2$ dBm, and the maximum Q-factor is improved by $\sim0.7$ dB. Our results confirmed that the phase conjugation property of the idler contributes to mitigation of nonlinear phase noise in addition to amplitude doubling. Compared with conven-
tional transmission without DSC/PSB, the PSB scheme achieves an improvement of 2.4 dB at the optimum Q-factor.

Finally, with the launch power set to the optimal values (4.2 dBm for PSB and 5.2 dBm for DSC), the number of loops is swept from 2 to 16 for investigating the system reach at a hard-decision forward error-correction (HD-FEC) limit of 8.5 dB. The results in Figure 5.4(d) show that DSC increases system reach from 10 loops (4,800 km) to 14 loops (6,720 km), while PSB increases reach further to 16 loops (7,680 km), representing a 60% improvement over conventional transmission. For multi-channel PSB, the optical pump should be placed at a proper wavelength, for instance on one side of all signal channels, so that the newly generated idlers do not overlap with original signals. Since the signal channel power is relatively low, the nonlinear effect
between signal channels is negligible during optical phase conjugation. Considering the different spectral spaces between signal channels and the optical pump, FWM efficiency may vary for each signal channel. Thus channel power equalization for idler channels might be needed after optical phase conjugation.

5.5 Simulation comparison between PSA system and EDFA system with and without PSB

Recent advances in non-degenerate PSA [107][108][109], which has potential NF improvement compared to EDFA and is modulation format independent, have revived interest in PSA for long-haul applications. Non-degenerate PSA relies on the simultaneous transmission of two correlated signal bands; the original signal and its phase conjugated idler, along with the pump, to allow phase-sensitive FWM interaction at the repeater, in order to suppress the uncorrelated noise in the two signal bands through constructive interference. In theory, a non-degenerate PSA can outperform an EDFA by as much as 6 dB in NF if only one signal band is transmitted for the EDFA system. If two signal correlated bands are used, just like a PSA link, for an EDFA link using a method such as PSB, the NF improvement of the PSA will be 3 dB, which is still an impressive gain and equivalent to doubling the transmission distance if all other transmission parameters are equal. However, in order to perform “phase-sensitive” amplification at the repeater, a total phase matching condition is required for the signal, idler, and the pump. Assuming advances in optical phase-locked loops technology can perfectly recover the correct pump phase for FWM, perfect dispersion compensation is required after each fiber span to preserve the phase coherence between the signal and idler. In long-haul transmission, it is well known that an in-line dispersion compensated (DC) link, like the one required for PSA operation, has much larger NL penalty than a dispersion uncompensated (DU) link due to the
fact that nonlinear interactions between neighboring symbols within a channel or between adjacent channels that cannot be effectively averaged out by fiber dispersion. Therefore, the optimal signal transmission power in a DC link has to be lower than a DU link to avoid higher nonlinear penalty. Before considering PSA for long-haul transmission, it is paramount that the fiber nonlinear impact due to in-line DC also be investigated. In this section, we conducted simulation study comparing the transmission performance of a DC PSA link with a DU EDFA link, using WDM 128-Gb/s DP-QPSK signals. We also applied our proposed PSB scheme in the DU EDFA link to observe its improvement compared to the DC PSA link under the same spectral efficiency condition. For ease of comparison, we considered ideal operating conditions that are achievable in theory for both PSA and EDFA in terms of amplification performance. This is the first time, to the best our knowledge, that PSA and EDFA are investigated together and fairly compared for practical long-haul applications.

5.5.1 Analytical Modeling

The configurations for the DC-PSA link and DU-EDFA link we used for the simulation are shown in Figure 5.5. VPITransmissionMaker™ is employed as the platform for the numerical simulation of the fiber transmission, based on the split-step Fourier method [110] for solving the nonlinear Schrödinger equation. For the DC-PSA link, the noise source will be the quantum noise amplified and suppressed by the PSA. For the DU-EDFA link, the main noise contribution will be from ASE noise added by the optical amplifiers. Since for long-haul transmission, signal performance is dominated mainly by the accumulated in-line noise, the receiver noises such as thermal noise and shot noise are neglected. We also assume that very high OSNR (>30 dB) can be achieved at the output of the transmitter, such that the noise added at the transmitter side before launching into transmission links can be omitted.
The signal generation and detection schemes for the WDM 128-Gb/s DP-QPSK channels are similar to the setups described in the earlier sections. To generate the signal and idler using an optical conjugate copier (see Figure 5.1) for the DC-PSA link, we place the signal and the pump wavelengths at 1550.12 nm and 1550.32 nm respectively, so the idler signal will be located at 1550.52 nm with 50-GHz spacing between the signal and idler. Optical equalization was performed to equalize the powers of the signal and the idler before they are transmitted together with other channels.

Both linear and nonlinear effects in the transmission fiber are modeled in our simulation. The linear impairments include fiber attenuation, CD, and birefringence-induced PMD. In digital coherent communication systems with adaptive equalizing filters implemented in the receiver DSP, fiber PMD will not cause additional penalty. For PSA systems however, PMD will be an issue because of the stringent phase-matching requirement and will have to be addressed either by designing fiber with extremely-low PMD or inserting an adaptive optical equalizer for realistic PSA implementation. In our simulation, fiber PMD is set to zero as we assumed the ideal phase-matching condition for PSA. For nonlinear impairments, only Kerr induced...
nonlinear effects, such as SPM, XPM and FWM, are considered as they dominate over other nonlinear effects in long-haul WDM systems. The evolution of optical signals in the transmission fibers, including the linear and nonlinear effects, is obtained by solving the nonlinear Schrödinger equation, which is modeled by the VPITransmissionMaker™ fiber toolbox. In the simulation, both the DC-PSA link and the DU-EDFA link use 80-km long single mode fiber (SMF) for each span, and the fiber parameters are as shown in Table 5.1.

<table>
<thead>
<tr>
<th>Fiber parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Span length</td>
<td>80 km</td>
</tr>
<tr>
<td>Attenuation</td>
<td>0.2 dB/km</td>
</tr>
<tr>
<td>Dispersion</td>
<td>16 ps/nm-km</td>
</tr>
<tr>
<td>Dispersion slope</td>
<td>0.08 ps/nm²-km</td>
</tr>
<tr>
<td>Kerr nonlinear index</td>
<td>$2.6 \times 10^{-20}$ m²/W</td>
</tr>
<tr>
<td>Effective area</td>
<td>80 µm²</td>
</tr>
<tr>
<td>PMD</td>
<td>0 ps/km¹/²</td>
</tr>
</tbody>
</table>

In the DU-EDFA link, OSNR is reduced after each span by the ASE noise contributed by each EDFA, and consequently caused the system performance degradation. In our simulation, ASE noise is described by the additive white Gaussian noise (AWGN) model and added to the optical signal at each EDFA. The NF of the EDFA is well known and given by:

$$NF_{EDFA} = \frac{2n_{sp}(G - 1)}{G} \sim 2n_{sp}$$

(5.3)

where $G$ and $n_{sp}$ are the amplifier gain and the spontaneous emission factor respectively. For comparison purposes, we consider the ideal case for EDFA operation with $n_{sp} = 1$ and an effective 3-dB NF.

To model the non-degenerate PSA for the WDM 128-Gb/s DP-QPSK signals which are modulated in both polarizations, a polarization diversity scheme has to be used for amplifying signals in two polarizations, as shown in Figure 5.6(a). The polar-
ization rotation and signal dispersion caused by the transmission fiber are assumed to be perfectly compensated before the PSA using polarization tracking and dispersion compensation modules with no extra insertion loss. Typically, the original pump signal from the optical conjugate copier at the transmitter is transmitted together with the signal and idler. To generate the pump for coherent FWM in the PSA, it will be filtered, amplified, and regenerated using an optical phase-locked loop to achieve the phase-matching condition. In our simulation, we assumed the pump regeneration is done perfectly at each repeater so we simply combine a pump source, which has the same phase as the original pump, with the signal using a zero-loss WDM multiplexer at 45° polarization relative to the signal. The linewidth of all the pump laser sources, which are all regenerated from the same pump laser used for the optical conjugate copier, are set to 100 kHz, a typical value for external cavity lasers (ECL) commonly used in digital coherent systems. All other noise from the pump source is assumed to be completely removed from the pump regeneration process and therefore neglected in our simulation, so no additional noise from the pump will be transferred to the amplified output signals. After using a PBS to perfectly separate the X- and Y-polarizations, the signal, the idler, and the pump in each polarization will undergo FWM in a piece of 200-meter-long HNLF to achieve non-degenerate phase-sensitive amplification for the signal and idler. Since we are interested in obtaining the optimal theoretical performance limit of the non-degenerate PSA, nonlinear processes such as SBS and stimulate Raman scattering (SRS) which are detrimental to PSA operation, are not taken into account in the simulation. To maximize the coherent gain and flatness of the gain wavebands of the PSAs, the parameters for the HNLF, as shown in Table 5.2, are fine-tuned to achieve the phase-matching condition across 100 GHz at both signal and idler bands. After FWM, the amplified signal and idler pair in two polarizations are combined using another PBS. The pump and unwanted spectral
by-products of the PSA are filtered out before the signals are launched into the next fiber span.

Figure 5.6: Simulation models for (a) non-degenerate PSA with polarization diversity and (b) quantum noise from fiber attenuation. $E_{in}$: input optical field into the fiber; $E_{out}$ output optical field from the fiber; $n_{qn}$ quantum noise.

Table 5.2: Parameters for highly nonlinear fiber used in the simulation

<table>
<thead>
<tr>
<th>HNLF parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>200 m</td>
</tr>
<tr>
<td>Attenuation</td>
<td>0 dB/km</td>
</tr>
<tr>
<td>Dispersion</td>
<td>−40 to 40 ps/nm-km</td>
</tr>
<tr>
<td>Nonlinear parameter ($\gamma$)</td>
<td>$23.58 \times 10^{-3} \text{m}^{-1}/\text{W}$</td>
</tr>
<tr>
<td>Effective area</td>
<td>40 $\mu$m$^2$</td>
</tr>
<tr>
<td>PMD</td>
<td>0 ps/km$^{1/2}$</td>
</tr>
</tbody>
</table>

For noise modeling in the PSA link, we have to take into account the quantum noise induced by the vacuum fluctuations. The noise, which is caused by the quantization of the optical field and described by Heisenberg’s uncertainty principle, can be modeled as an additive noise with constant spectral density in both orthogonal
polarization states:

\[ \langle |n_{\text{qn}}| \rangle = h\nu/2 \quad (5.4) \]

where \( h \) is Planck's constant and \( \nu \) is the optical frequency. For DU-EDFA links, the quantum noise is already included within the ASE noise in the EDFA model. However, for DC-PSA simulation, quantum noise has to be added into the link as the VPITransmissionMaker\textsuperscript{TM} fiber models do not include quantum noise due to fiber attenuation. A semi-classical model is used in our simulation to emulate the coupling of the signal and quantum noise \[113\], as shown in Figure 5.6(b). The optical field at the end of span is given as:

\[ E_{\text{out}} = \sqrt{\Gamma} E_{\text{in}} + \sqrt{1 - \Gamma} n_{\text{qn}} \quad (5.5) \]

where the splitting ratio \( \Gamma \) is determined by the span attenuation, and \( E_{\text{in}} \) is the input optical field launched into the fiber span. In an ideal PSA, the input additive quantum noise on both the signal and idler is amplified and summed. Because the newly added quantum noises in the signal and idler bands after each fiber span are uncorrelated, the ideal PSA operation will provide 6-dB additional gain (four-fold power gain) for coherent signal-idler pair over the uncorrelated noise. The noise figure of PSA can then be expressed as:

\[ \text{NF}_{\text{PSA}} = \frac{P_{\text{in}}}{G P_{\text{in}}} \frac{\Delta f}{2 \frac{h \nu}{2} \Delta f} = \frac{1}{2} \quad (5.6) \]

where \( P_{\text{in}} \) is signal input power to the amplifier, \( G \) is the signal gain, and \( \Delta f \) is the signal bandwidth. Comparing Equation 5.3 and Equation 5.6, we observe that the ideal PSA with a signal-idler pair has a 6-dB improvement in NF compared to an ideal EDFA. The theoretical PSA gain difference between the signal and the noise, when only signal band is considered, is 3 dB. In Table 5.3, we compare the
gain values achieved in the designed PSA by using either theoretical calculation or VPITransmissionMaker™ simulation. We find that even though the achieved gain value is ~0.3 dB less than what can be achieved in theory, probably due to slightly imperfect phase-matching, the 2.9-dB gain benefit between the signal and noise is very close to the theoretical value. This enables transmission simulations using the designed PSAs in close to optimal conditions for comparison purpose.

Table 5.3: Performance of the designed PSA: theory and simulation

<table>
<thead>
<tr>
<th></th>
<th>Theoretical</th>
<th>Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-sided Gain (signal)</td>
<td>17.46 dB</td>
<td>17.03 dB</td>
</tr>
<tr>
<td>Gain (noise)</td>
<td>14.46 dB</td>
<td>14.13 dB</td>
</tr>
<tr>
<td>Gain benefit</td>
<td>3 dB</td>
<td>2.9 dB</td>
</tr>
</tbody>
</table>

5.5.2 Simulation results and discussions

We first compare the results of single channel transmission over the DC-PSA link and DU-EDFA link. Even in single channel transmission, both the signal and idler have to be transmitted over the DC-PSA link, occupying twice the optical bandwidth as in the DU-EDFA link. A total of 140 SMF spans of 80-km length, the equivalent of 11,200-km transmission distance, is used for our simulation. The received signal Q-factor is plotted versus the channel launch power per span in Figure 5.7. For the DU-EDFA link, the optimal launch power is +1 dBm with Q value of 11.1 dB. By switching to the DC-PSA link, the received signal Q-factor can have as much as 7-dB gain in the linear regime due to much reduced NF of PSA and increased signal power from combining the signal and idler pair. However, because of the requirement of complete dispersion compensation at each span, the signal will experience a much more severe fiber nonlinear effect which significantly lowers the optimal channel launch power to ~−3 dBm. The received Q-factor is 15.2 dB, more than 4 dB higher compared to the DU-EDFA link. However, when we apply the PSB scheme on the DU-EDFA link
using the same optical bandwidth as the DC-PSA link, the Q-factor can be raised to 13.5 dB, only 1.7 dB lower than the performance of the DC-PSA link.

Figure 5.7: Received signal Q-factor vs. channel launch power for different transmission schemes in single channel operation.

WDM transmission of eight 127-Gb/s DP-QPSK signals with 50-GHz channel spacing for both links is performed with the results plotted in Figure 5.8. For the DC-PSA link, the signal and idler pair is placed at the center of the transmitted spectrum. The same channel configuration is also used for the PSB scheme in the DU-EDFA link. The performance of the DC-PSA link degrade further due to inter-channel nonlinearty such as XPM and FWM. The optimal channel launch power is reduced to −5 dBm with a Q-factor of 13.7 dB. On the other hand, the additional nonlinearity impact due to WDM transmission is not as significant in the DU-EDFA link, with only 0.6-dB drop to 10.5 dB. With the PSB scheme applied, the Q-factor can be improved up to 13.3 dB, which is only 0.4 dB lower than in the DC-PSA
scheme. Our results show that even under the ideal operating condition, PSA still faces the challenge of higher fiber nonlinearity impact due to its requirement for perfect span-by-span dispersion compensation. Therefore, the much sought-after OSNR gain due to NF reduction will be almost entirely wasted because of the high fiber nonlinearity impact. Consequently, other than the implementation complexity, further efforts in nonlinearity mitigation will be required for PSA-based links before they can be suitable for long-haul systems. Meanwhile, schemes such as PSB which utilize transmission of phase conjugated copies in EDFA links can achieve almost the same performance as PSA-based links, by enhancing signal tolerance to OSNR degradation and fiber nonlinearity.

Figure 5.8: Received signal Q-factor vs. channel launch power for different transmission schemes in WDM operation.
5.6 Conclusion

We demonstrated a novel phase-sensitive boosting scheme to improve noise and non-linearity tolerance in long-haul WDM transmission. The scheme utilizes copropagation of a phase-conjugated idler signal, which is generated by FWM and transmitted together with the original signal. After transmission, the signal and idler are jointly detected and processed in DSP to suppress nonlinear phase distortion and increase OSNR. The proposed digital PSB scheme is independent on modulation format and does not require an optical phase-locked loop to achieve phase matching required by conventional PSAs. The numerical simulations verify the principle of the proposed PSB scheme, and show 2.5-dB Q-improvement after 6,300-km transmission over both dispersion-managed and dispersion-unmanaged links. Our experimental result shows that the PSB scheme achieves a 2.4-dB Q-improvement over conventional transmission after 4,800 km of DMF at the expense of 50% reduction in spectral efficiency, while system reach using HD-FEC can be increased by 60% to 7,680 km. Compared to a PSA-based link, the PSB scheme does not require in-line dispersion compensation which can significantly increase the nonlinearity penalty on transmission performance. In addition, we compared the performance of a DC-PSA link and a DU-EDFA link by numerical simulation, and found that by using PSB, we can achieve almost the same transmission performance as the PSA-based link, even though PSA has 6-dB lower NF than EDFA. Compared to the complexity and challenges of implementing PSA, our proposed PSB scheme is much more cost-effective while delivering comparable performance in long-haul transmission.
Chapter 6

Conclusions

With the fast development of optical and electro-optical devices, all-optical signal processing of high-speed light signals, enabled by nonlinear optical components, has attracted increasing interest in the areas of telecommunications, computing and military applications, while its electric counterpart is approaching physical limits that are hard to exceed. Thanks to the intrinsic merits of vast bandwidth resources, extremely high capacity, low crosstalk etc., ultrafast all-optical signal processing enabled by nonlinear optics is considered as a promising alternative for future signal processing and computing. Recent progress in material manufacturing, packaging and integration results in high performance, low cost and compact nonlinear devices. These advanced nonlinear devices have led to the reduction of the power required for nonlinear optical signal processing down to levels compatible with compact semiconductor lasers, and have paved the way to on-chip integration, which is bringing nonlinear optical signal processing to practical deployment.

In this context, high-performance neuromorphic signal processing and computing systems inspired by biological neural networks can benefit from nonlinear optical processing technologies. Using the high-speed and broadband photonic devices, neuromorphic signal processing and computing platforms could solve unconventional
computing problems and outperform current technology in both power efficiency and complexity. Wedding the neuro-ethology drawn from biological neurons with sophisticated modern engineering techniques not only helps in studying biological neural circuits, but also opens up a wide range of applications such as adaptive control, learning, perception, motion control, sensory processing (vision systems, auditory processors, and olfactory systems), and autonomous robots. For example, using analog very large scale integration technology (VLSI), small low-power front end sensor devices have been implemented that closely replicate the capabilities of the retina and the cochlea.

On the cellular level, neurons operate on information encoded as “spikes”, a type of signal with both analog and digital properties. Spike processing exploits the efficiency of analog signals while overcoming the problem of noise accumulation inherent in analog computation. The basic biological structure of a spiking neuron, or leaky-integrate-and-fire (LIF) neuron, consists of a dendritic tree that collects and sums inputs from other neurons, a soma that acts as a low pass filter and integrates the signals over time, and an axon that carries an action potential, or spike, when the integrated signal exceeds a threshold. Neurons are connected to each other via synapses, or extracellular gaps, across which chemical signals are transmitted. The axon, dendrite, and synapse all play an important role in the weighting and delaying of spike signals.

In our lab, the first photonic neuron bench-top prototype was previously presented and demonstrated to emulate an LIF neuron using ultrafast photonic devices. The temporal integration of the input optical pulses was completed by a semiconductor optical amplifier (SOA) -based integrator. A following nonlinear optical loop mirror (NOLM) with a Ge-doped fiber was employed as an ultrafast optical thresholder. Since then, efforts have been made to improve the photonic neuron’s elemental designs. With the help of new techniques and components, the spiking neuron’s per-
formance and functionality have been significantly enhanced. To enable the temporal integrator to integrate both excitatory and inhibitory inputs, we studied the gain pumping effect in a C-band SOA and its simultaneous operation with gain depletion. The experimental results showed the feasibility of using a C-band SOA as a versatile integrator for both types of inputs. Another scheme by using cross-absorption modulation (XAM) in an electro-absorption modulator (EAM) was also proposed and demonstrated to achieve non-inverted integration, based on which an asynchronous spiking neuron was demonstrated. Thus truly asynchronous processing is realized in our photonic neuron, which guarantees high processing efficiency and capacity, and eliminates the possibility of triggering multiple pulses by one input. In addition, a novel optical thresholder based on a Ge-doped NOLM with band-pass filter (BPF) configuration is proposed to greatly improve the power transfer function, in terms of the steepness of the slope, as well as the flatness of the high and low output levels. Furthermore, we proposed two neuromorphic circuits with our enhanced photonic neurons, i.e. a signal feature recognition device inspired by the escape response of a crayfish and a timing jitter insensitive optical logic gate reconfigurable between AND and NOT operation.

To realize the most powerful capability of a biological neuron—learning, an optical spike timing dependent plasticity (STDP) device is built by utilizing the inherent carrier dynamics in an SOA and an EAM. STDP modifies synaptic strengths based on the correlation between the pre- and post-synaptic activities, allowing the neuron to learn according to the environment or a teacher. With the optical STDP device, we successfully demonstrated, for the first time, supervised learning in an enhanced photonic neuron, which forms the foundation for a photonic neuron to learn about the environment like a biological brain, while having a processing speed that is a billion times faster.
Other than photonic neurons, we also investigated another area in need of ultra-fast signal processing technology—telecommunications. As the sophisticated coherent detection and digital signal processing (DSP) technologies have pushed the single channel transmission capacity to more than 1 Tb/s in long-haul wavelength-division multiplexing (WDM) transmission, it is extremely difficult for electronic devices to accomplish real-time processing at such a high operational speed. Therefore we proposed and demonstrated a scheme using all-optical phase conjugation to mitigate the nonlinear impairments in coherent ultra-long-haul transmission systems. The scheme utilizes a phase-conjugated idler signal, which is generated by ultra-fast four-wave mixing (FWM) and transmitted together with the original signal. After transmission, the signal and idler are jointly detected and processed to suppress nonlinear phase distortions due to long-haul transmission and increase optical signal-to-noise ratio (OSNR). Both numerical simulations and experiment demonstrations show a significant improvement of signal quality (2.4 dB in terms of Q-factor) and large increase of system reach (60%).

Although our bench-top spiking neuron prototypes based on discrete photonic components are bulky, complex, power-hungry and hard to scale up to more than a few neurons, with the fast progress of nonlinear materials, manufacturing and integration, it is promising to have our bench-top prototypes integrated onto a single chip. Upon large-scale integration and interconnection of these photonic neurons, more complex tasks such as unconventional computing and learning akin to functions in human brains may be achieved.
Bibliography


