

Photonic Technologies for Ultrafast Neuromorphic Information Processing

Dafydd Owen-Newns



Institute of Photonics

University of Strathclyde

January 2026

Thesis presented in fulfilment of the requirement for the
degree of Doctor of Philosophy

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree. The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Abstract

The rapid growth of artificial intelligence has placed increasing demands on digital computing hardware, motivating the exploration of alternative computing paradigms that offer improved efficiency, bandwidth, and scalability. Neuromorphic computing, inspired by the event-based and energy-efficient operation of biological neural systems, represents one such paradigm. In parallel, photonic technologies offer unique advantages for neuromorphic processing, including ultrafast signal propagation, high bandwidth, low communication losses, and intrinsic parallelism. This thesis investigates the use of two photonic and optoelectronic devices as photonic spiking neurons for neuromorphic computing.

The first part of the thesis focuses on Vertical-Cavity Surface-Emitting Lasers (VCSELs) operated under optical injection. It is shown experimentally that VCSELs can exhibit rich nonlinear dynamics that closely resemble key neuron-like behaviours, including excitability, thresholding, temporal integration, refractoriness, and spike generation on sub-nanosecond timescales. These properties are exploited to realise photonic spiking neural networks using time-division multiplexing, enabling networks with large effective neuron counts operating at gigahertz rates. The computational capabilities of these systems are demonstrated through tasks such as nonlinear classification and chaotic time-series prediction, supported by training strategies adapted to the binary and temporal nature of photonic spikes.

The second part of the thesis investigates Resonant Tunnelling Diodes (RTDs) as compact optoelectronic spiking neurons. A physics-based circuit model is used to describe the excitable dynamics arising from the negative differential resistance of RTDs,

and numerical simulations are used to explore their spiking behaviour under optical and electrical perturbations. Experimental demonstrations validate these models and show that RTD neurons can operate as ultrafast, multi-modal spiking neurons. The applicability of RTD-based neuromorphic systems is illustrated through several tasks, including photonic spiking extreme learning machines for classification, recurrent architectures for tunable photonic memory, and event-based time-series processing. In particular, a dual optical–electrical modulation scheme is shown to enable efficient rising-edge detection in complex signals, bridging sensing and computation at the device level.

Acknowledgements

I would like to extend my deep thanks to my supervisor Prof. Antonio Hurtado for guiding me through the process of completing this PhD project. He has been persistently motivating, patient, and supportive, and always encouraging me to learn and drive my work forward. I also acknowledge the support of Fraunhofer UK, for sponsoring part of my project, and Dr. Henry Bookey for coordinating.

I would also like to acknowledge my fellow research group members, particularly Dr. Joshua Robertson and Dr. Matěj Hejda, for introducing me into the group and teaching me how to use the lab, present my results, and generally progress in my research. Also to Dr. Giovanni Donati, Dr. Maria Duque, Andrew Adair and Dylan Black, for sharing thoughts and conversations through the years. Thanks also go to the staff in the Institute of Photonics, particularly Sharon Kelly, Lorraine Annand and Leanore Ferrans, for always keeping the department running effectively and being able to solve any problem that arose.

My thanks and appreciation also go to my family and friends across the world for supporting me throughout and keeping me happy.

List of Publications

First Authored Publications

- D. Owen-Newns, J. Robertson, M. Hejda, and A. Hurtado, “Ghz rate neuromorphic photonic spiking neural network with a single vertical-cavity surface-emitting laser (VCSEL),” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 29, no. 2: Optical Computing, p. 1500110, 2023.
- D. Owen-Newns, J. Robertson, M. Hejda, and A. Hurtado, “Photonic spiking neural networks with highly efficient training protocols for ultrafast neuromorphic computing systems,” *Intelligent Computing*, vol. 2, p. 0031, 2023.
- D. Owen-Newns, L. Jaurigue, J. Robertson, A. Adair, J. Jaurigue, K. Lüdge, and A. Hurtado “Photonic spiking neural network built with a single vcsel for high-speed time-series prediction,” *Communications physics*, vol. 8, no. 1, pp. 110–9, 2025.
- D. Owen-Newns, J. Robertson, G. Donati, Q. Al-Taai, J. Figueiredo, E. Wasige, K. Lüdge, B. Romeira, A. Hurtado, ”Neuromorphic Photonic Processing and Memory with Spiking Resonant Tunnelling Diode Neurons and Neural Networks,” *Advanced Intelligent Systems*, p. 202500800, 2026

Co-Authored Publications

- G. Donati, et al. “Spiking Rate and Latency Encoding with Resonant Tunneling Diode Neuron Circuits and Design Influences.” *Physical Review Applied*, vol. 24, no. 2, 2025.

Chapter 0. List of Publications

- M. Hejda, et al., “Optical spike amplitude weighting and neuromimetic rate coding using a joint VCSEL-MRR neuromorphic photonic system,” *Neuromorphic Computing and Engineering*, vol. 4, no. 2, p. 24011, 2024.
- G. Donati, et al., “Spiking flip-flop memory in resonant tunneling diode neurons,” *Phys. Rev. Lett.*, vol. 133, p. 267301, 2024.
- M. Hejda, et al. “Programmable Optical Synaptic Linking of Neuromorphic Photonic-Electronic RTD Spiking Circuits.” *ACS Photonics*, vol. 11, no. 10, pp. 4279–87, 2024.
- W. Zhang, et al., “Photonic-electronic spiking neuron with multi-modal and multi-wavelength excitatory and inhibitory operation for high-speed neuromorphic sensing and computing,” *Neuromorphic Computing and Engineering*, vol. 4, no. 4, p. 44006, 2024
- Q. R. A. Al-Taai et al., “Optically-triggered deterministic spiking regimes in nanostructure resonant tunnelling diode-photodetectors,” *Neuromorphic Computing and Engineering*, vol. 3, no. 3, p. 34012, 2023.
- M. Hejda, et al. “Artificial Optoelectronic Spiking Neuron Based on a Resonant Tunnelling Diode Coupled to a Vertical Cavity Surface Emitting Laser.” *Nanophotonics* (Berlin, Germany), vol. 12, no. 5, pp. 857–67, 2023
- A. Skalli, et al. “Photonic Neuromorphic Computing Using Vertical Cavity Semiconductor Lasers.” *Optical Materials Express*, vol. 12, no. 6, pp. 2395–414, 2022

Contents

Abstract	ii
Acknowledgements	iv
List of Publications	v
List of Figures	xiii
List of Tables	xiv
1 Introduction	1
1.1 Artificial Intelligence	1
1.2 Neuromorphic Processing	6
1.3 Photonics as a Platform for Neuromorphic Processing	14
1.4 Thesis Outline	20
2 Spiking VCSEL-Based Photonic Neurons	21
2.1 Vertical Cavity Surface Emitting Lasers	22
2.1.1 Optical Injection Locking of VCSELs	23
2.2 Optical Spiking VCSEL Neurons	25
2.2.1 Experimental Setup	25
2.2.2 Excitable Optical Spikes	29
2.2.3 Spike Modulation	38
2.3 Summary	39

3	Photonic Neural Networks Based on VCSELs	41
3.1	VCSEL-based Photonic Spiking Neural Network	42
3.1.1	Nonlinear Classification Task Operation	44
3.1.2	Experimental Performance of the Photonic Spiking Neural Network	46
3.2	A Novel Training Method for the Spiking VCSEL p-SNN	56
3.2.1	Experimental Procedure	57
3.2.2	Ordinary Least Squares Training Method	61
3.2.3	Binary Weight Training	62
3.3	Time-Series Prediction using a VCSEL p-SNN	69
3.3.1	Experimental Results on the time-series Prediction Task	70
3.4	Summary	80
4	Resonant Tunnelling Diode Neurons	81
4.1	The Physics of Resonant Tunnelling Diodes	82
4.1.1	Resonant Tunnelling Diode Circuits	84
4.1.2	Current-Voltage characteristics of RTDs	85
4.1.3	RTD Circuit Dynamics	86
4.1.4	Optically Coupled Resonant Tunnelling Diodes	93
4.1.5	Numerical Simulation of RTDs	95
4.2	Spiking Neurons Based on Resonant Tunnelling Diodes	97
4.2.1	Thresholding	98
4.2.2	Refractoriness	101
4.2.3	Integrate and Fire	101
4.2.4	Resonate and Fire	103
4.3	Experimental investigation of RTD neurons	104
4.4	Summary	108
5	Neuromorphic Processing Using RTDs	110
5.1	Neuromorphic Processing with Single Photonic-Electronic RTD Neurons	111
5.1.1	Edge-Feature Detection in Time-Series	111
5.1.2	Spike Rate Encoding with RTD Neurons	119

Contents

5.2	Neuromorphic Processing with Arrays of RTD Neurons	120
5.2.1	Photonic Spiking Neural Network with uncoupled RTD neuron arrays	122
5.2.2	Neuromorphic Photonic Memory built with coupled RTD neurons	126
5.3	Summary	131
6	Conclusions	132
6.1	Conclusions	132
6.2	Future Work	138
6.2.1	Experimental and Theoretical Developments	138
6.2.2	Potential Applications of Neuromorphic Photonics	139
	References	141

List of Figures

1.1.1 Comparison of single perceptron and multi-layer perceptron	3
1.1.2 Construction of a function by summing activation functions	4
1.1.3 Sample of Neural Network architectures	6
1.2.1 Spike-Timing-Dependent-Plasticity	11
1.2.2 Leaky Integrate and Fire circuit	12
1.2.3 Response of a Resonate and Fire compared to Integrate and Fire neuron	13
1.2.4 Trajectories of a Hopf Bifurcation	15
2.1.1 Schematic of the structure of a VCSEL	23
2.1.2 Spectra of VCSELs at 1300 <i>nm</i> and 1550 <i>nm</i>	24
2.2.1 Experimental Setup used to implement spiking in a VCSEL	26
2.2.2 Photograph of the VCSEL platform	27
2.2.3 Photodetector trace of an optical spike emitted by a VCSEL	30
2.2.4 Traces of the orthogonal polarisation components of the VCSEL during a spike	32
2.2.5 Threshold and fire behaviour in a VCSEL neuron	34
2.2.6 Integrate and fire behaviour in a VCSEL neuron	35
2.2.7 Demonstration of the refractory period of a VCSEL neuron	37
2.2.8 Experimental setup used to modulate the output of a VCSEL neuron . .	39
2.2.9 Modulating spikes from a VCSEL neuron	40
3.1.1 Experimental setup and operation of a VCSEL p-SNN	47
3.1.2 Scatter plot of the Iris dataset	48

List of Figures

3.1.3 Logical architecture of the VCSEL p-SNN	49
3.1.4 Input and Output time-series of the VCSEL p-SNN for a sample of the Iris data set	50
3.1.5 Spike patterns produced by the VCSEL p-SNN for the Iris data set . . .	52
3.1.6 Confusion matrix produced by the VCSEL p-SNN for the Iris data set .	54
3.1.7 Classification error of the VCSEL p-SNN on the Iris dataset against training set size	55
3.2.1 Input and output traces of the VCSEL p-SNN for a sample of the MADE- LON data set	58
3.2.2 Spike patterns produced by the VCSEL p-SNN for the MADELON data set	60
3.2.3 Confusion matrices produced by the VCSEL p-SNN for the MADELON data set	62
3.2.4 Demonstration of significance scoring of spiking nodes	65
3.2.5 Confusion matrices produced by the VCSEL p-SNN for the MADELON data set using significance training	66
3.2.6 Classification accuracy of the VCSEL p-SNN on the MADELON data set using significance training and 2048 nodes	67
3.2.7 Classification accuracy of the VCSEL p-SNN on the MADELON data set using significance training and 4096 nodes	68
3.3.1 Layout and architecture of the VCSEL p-SNN used for time-series pre- diction	70
3.3.2 A sample of the Mackey-Glass time-series	72
3.3.3 Operation of the VCSEL p-SNN for tiem series prediction	72
3.3.4 Input and output traces of the VCSEL p-SNN for time-series prediction	74
3.3.5 Spike patterns generated by the VCSEL p-SNN using the Mackey-Glass time-series	74
3.3.6 Performance of the VCSEL p-SNN in multi-step prediction of the Mackey- Glass time-series	76

List of Figures

3.3.7 Performance against training set size of the VCSEL p-SNN on the prediction of the Mackey-Glass time-series	77
3.3.8 Examples of predictions of the Mackey-Glass time-series by the VCSEL p-SNN	79
4.1.1 Conduction band structure of a Resonant Tunnelling Diode	83
4.1.2 Lumped circuit model describing the RTD circuit	84
4.1.3 RTD IV curves for different parameter values	87
4.1.4 Phase plot of the nullclines of the RTD circuit dynamics	88
4.1.5 Phase plot of the orbits of the RTD circuit when biased at 0.65 V	90
4.1.6 Phases of relaxation oscillations in the RTD circuits	92
4.1.7 Traces of voltage, current, carrier count and laser output during an RTD-laser spike	96
4.2.1 Phase plot of trajectories of the RTD circuit returnign to equilibrium . .	99
4.2.2 Demonstration of the threshold and fire mechanism in an RTD neuron .	100
4.2.3 Threshold amplitude of optical pulses against bias voltage in an RTD neuron	101
4.2.4 Refractory period of spikes from an RTD neuron	102
4.2.5 Integrate and fire in an RTD neuron	103
4.2.6 The oscillating response of an RTD to small perturbations when biased in the valley	104
4.2.7 Resonate and fire in an RTD neuron	105
4.3.1 Experimental setup used to implement RTD neurons	107
4.3.2 Experimentally observed IV curve of an RTD under darkness and illumination	108
4.3.3 Trace of the voltage across an RTD during a spike	109
5.1.1 Varying optical trigger threshold by varying the bias in the IV valley . .	112
5.1.2 Operation of an event-based edge detection system based on a n RTD neuron	114

List of Figures

5.1.3 Simulation of the rising edge detection by an RTD on the Mackey-Glass time-series	115
5.1.4 Experimental results showing the rising edge detection by an RTD on a sample of the Mackey-Glass time-series	117
5.1.5 Edges in the MAckey-Glass time-series detected by an RTD neuron . . .	118
5.1.6 Variation of the constant spiking rate of and RTD with bias voltage inside the NDR	119
5.1.7 Simulation of spike-rate encoding of an 8 level signal	121
5.2.1 Architecture of a spatially multiplexed neural network of RTD neurons .	123
5.2.2 Spike patterns generated by the RTD array for the Iris data set	125
5.2.3 Single RTD based autaptic spiking memory	128
5.2.4 schematic and operation of a multiple-RTD based tuneable depth spiking memory system	129

List of Tables

1.1	Features to take into account when designing hardware-based spiking neural networks.	10
3.1	Error and correlation against training set size. Experimentally obtained values of NMSE (Normalised Mean Squared Error) and correlation coefficient for 10-steps ahead predictions for three different configured network node architectures in the p-SNN and different training set sizes (as a percentage of 8000).	80
4.1	Parameters used in the definition of the IV curve.	86
4.2	Values of resistance (R), capacitance (C) and inductance (L) used for simulating ultrafast RTD based neurons.	92
4.3	Parameters used to model a simple two-level diode laser.	95
4.4	Material composition of the photonic-electronic RTD device. This was designed and fabricated by our project collaborators at the University of Glasgow (Dr Qusay Al-Taai, Prof Edward Wasige). In bold are the layers that form the double-barrier quantum well, responsible for the nonlinear dynamics of the RTD.	106

Chapter 1

Introduction

This chapter introduces the background to this thesis and the motivation of the work that will be presented in later chapters: The basic principles of artificial neural networks will be explored, and their generalisation to spiking neural networks will be presented, motivating the development of neuromorphic computing. Neuromorphic processing, inspired by the computational power of the brain, will be described, particularly the dynamics of biological neurons. Photonics will be introduced as an alternative to electronic platforms, and the state-of-the-art technologies in the particular field of photonic neuromorphics will be explored.

1.1 Artificial Intelligence

”Artificial Intelligence” (or AI) has often been defined as the ability of computers to perform tasks associated with intelligence, such as decision making and pattern recognition, and being able to learn to do so autonomously. Originating as a field of study with Alan Turing’s work [1], and perhaps earlier with Ada Lovelace’s ideas of a ”musical composing machine” [2], the research into computed intelligence has gone through several cycles of intensive and quiet periods, until recent years when the development of graphics processing units allowed the widespread realisation of many complex machine learning systems.

Perhaps the most widely discussed concept in AI and machine learning is the Ar-

tificial Neural Network (ANN): This is a method of finding an approximation to a function, given a set of points generated by that function (i.e. the inputs and resulting outputs) by summing transformed copies of one particular function (known as the activation function)[3], generalising the notion of traditional curve fitting, which requires prior knowledge of the form of the function.

This powerful ability - to reconstruct the pattern that underlies a set of data, and then to make accurate predictions of further outputs - has found a range of applications, from general data processing, medicine, engineering, and more recently to image and text generative systems.

The first ANN to be created was the "perceptron", a simple binary classifier that separates data into one of two classes, by applying a linear transform to the data 'features', followed by a threshold operation to determine the classification[4]. This method is capable of discriminating classes of data that are separated by a plane in the feature space - and will not work if the separating surface between the classes is curved or not simply connected. The single perceptron can be extended to tackle these problems by combining the outputs of several parallel perceptrons - each with their own set of weights and biases - into the input to a further perceptron.

Perceptrons were later generalised by replacing the discontinuous threshold with smooth continuous activation functions. The result of this is that the output of the network can follow the value of a real-valued function of the input features, rather than a yes/no classification. These ANNs proved very successful and gained widespread applications due to the ability to approximate any arbitrary function given a sufficiently large number of nodes, known as the universal approximation theorem[3].

The output of a single hidden layer ANN can be written as:

$$\hat{\mathbf{f}}(x) = \mathbf{W}_{out}\sigma(\mathbf{W}_{in}\mathbf{x} - \mathbf{b}), \quad (1.1)$$

where $\mathbf{x}$ is the vector representing the input data point, $\mathbf{W}_{in/out}$ represent the linear transforms (weights) between layers (from input to hidden, and hidden to output), $\mathbf{b}$ is the vector of bias terms added to each node in the hidden layer, and σ is the element-

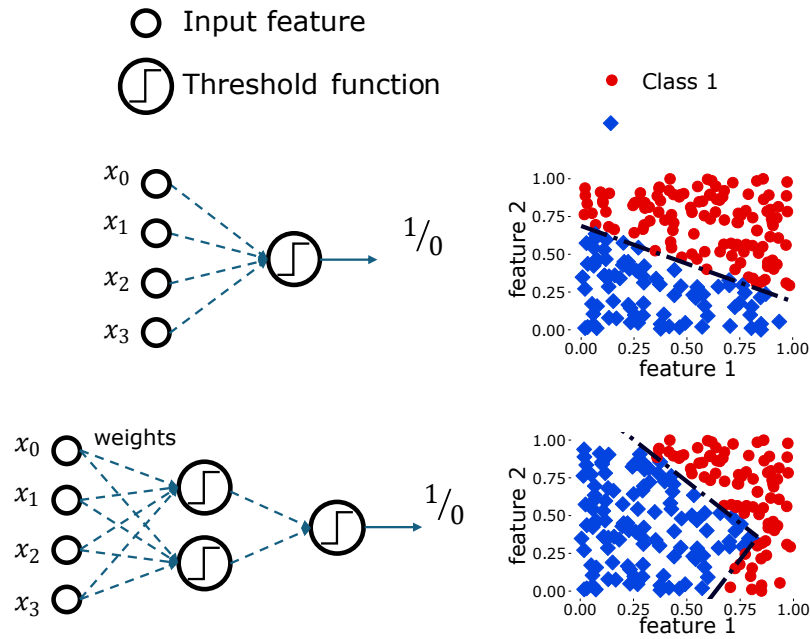


Figure 1.1.1: Classification by a single Perceptron vs. a multi-layer Perceptron. A single perceptron calculates the dot product of the features with a vector (the weights) and checks whether the result is greater than some value (the bias). This operation splits the feature space along a straight line. The multi-layer perceptron further combines the results of two different single perceptrons, thus allowing a classification boundary formed by joining two lines at their intersection. Increasing the number of thresholding nodes in the 'middle' layer allows for more independent pieces in the decision boundary.

wise nonlinear "activation function" that gives the output value of the nodes in the hidden layer of the network.

This represents the summation of several copies of the activation function, shifted and scaled by the input linear transformations. Figure 1.1.2 shows how this superposition forms the total output function, demonstrating the same set of weights (input and output linear transforms), for two commonly used activation functions, the sigmoid and rectified linear unit.

Extending these ANNs come deep neural networks (DNNs): approximating a function not only by superposing one set of activation functions, but by composing several 'layers' of ANNs sequentially. These deep neural networks have proven to be more powerful, yet require more complex training methods, since the calculations required to find an analytic solution to the optimisation problem (involving inversion of very large

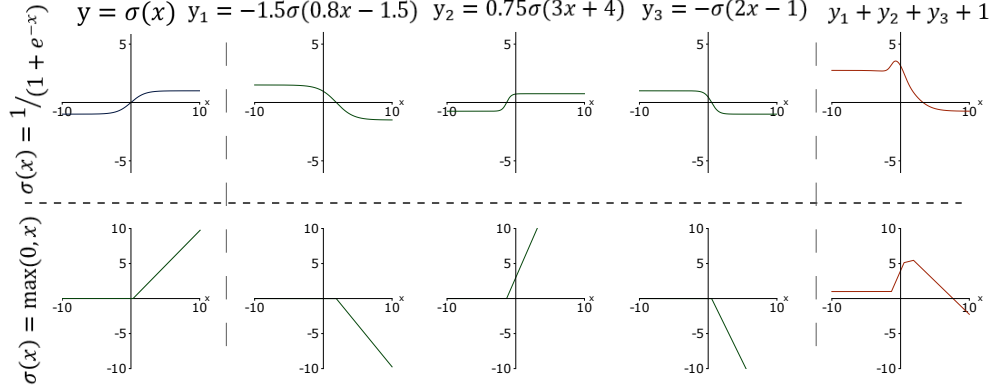


Figure 1.1.2: Construction of a nonlinear function by summation of an activation function after scaling and shifting - as is done in a neural network. Each row shows the result for a different activation function, with the same weights and biases applied. Top uses the logistic sigmoid function, while the bottom uses the rectified linear unit.

matrices) may be intractable. The process of backpropagation, based on the chain rule of differentiation and the Taylor expansion of the neural network's output[5–7] is used widely to find the gradients in the network used to calculate incremental updates to the network parameters during training. Variations on these DNNs, such as convolutional neural networks (CNNs) or physically-informed neural networks (PINNs), make use of specialist knowledge of the form of the data to instruct the particular architecture of the network, for example the translation-symmetric kernels of a CNN. These specialist architectures - as opposed to a fully connected deep neural network - allow for faster and more accurate training, reducing implementation costs. CNNs (and their descendants) in particular have proven very powerful for image processing[8], spurring the further development of the attention mechanism and transformer networks[9] that underlie the modern Large-Language-Models (LLMs).

Additional neural network architectures of note include the Hopfield network[10, 11], for which John Hopfield was awarded the 2024 Nobel prize[12], a form of recurrent neural network (RNN: these are DNNs that have recurrent connections - the output of nodes deeper in the network are fed into the input of earlier nodes) that acts as a system with "associative memory" that can learn input patterns of neuron activity through adjusting the weights between neurons in the network, and subsequently retrieve those stored patterns when presented with an incomplete input pattern; Ising

machines, which aim to solve certain problems by mapping their solutions to the Ising model of spin-glasses, and processing in a manner similar to that of Hopfield networks; and reservoir computers[13–15], designed with hardware implementation in mind, which are large-scale recurrent neural networks with fixed, randomly set internal weights and a trained output layer. These reservoir computers can be implemented physically by using a nonlinear medium (such as waves on the surface of a reservoir of water, hence the name) as the hidden, recurrent network of nodes, and mapping the read-out values from this medium to the output using a trainable linear transform. The benefit of using a reservoir computer is the significant reduction in training requirements for similar performance to deep neural networks - since only the final, output layer needs training, and thus can be trained using exact methods like single-layer perceptrons, and the potential for a relatively simple hardware implementation. Between reservoirs and feedforward networks (networks that explicitly contain no recurrent connections) lie an ANN architecture sometimes called "extreme learning machines"[16], which are two-layer feedforward ANNs in which the input layer weights are chosen randomly and fixed, while the output weights are trained. This has similar advantages to reservoir computers, able to be implemented on hardware that can apply the input linear transform and nonlinear activation, and the training of the output weights can be done by well-known, simple algorithms such as linear-least squares fitting, rather than relatively more costly backpropagation.

All this being said, the result is that any machine that is designed to run an artificial neural network (of almost any type) requires two key capabilities: *Linear Transformations*, to apply weights (typically done by matrix-vector multiplication); and a *nonlinear activation function*, which can be applied element-wise to the input data vectors (the choice of this function can greatly affect the performance and training speed of the network, and even determine whether the ANN will work at all).

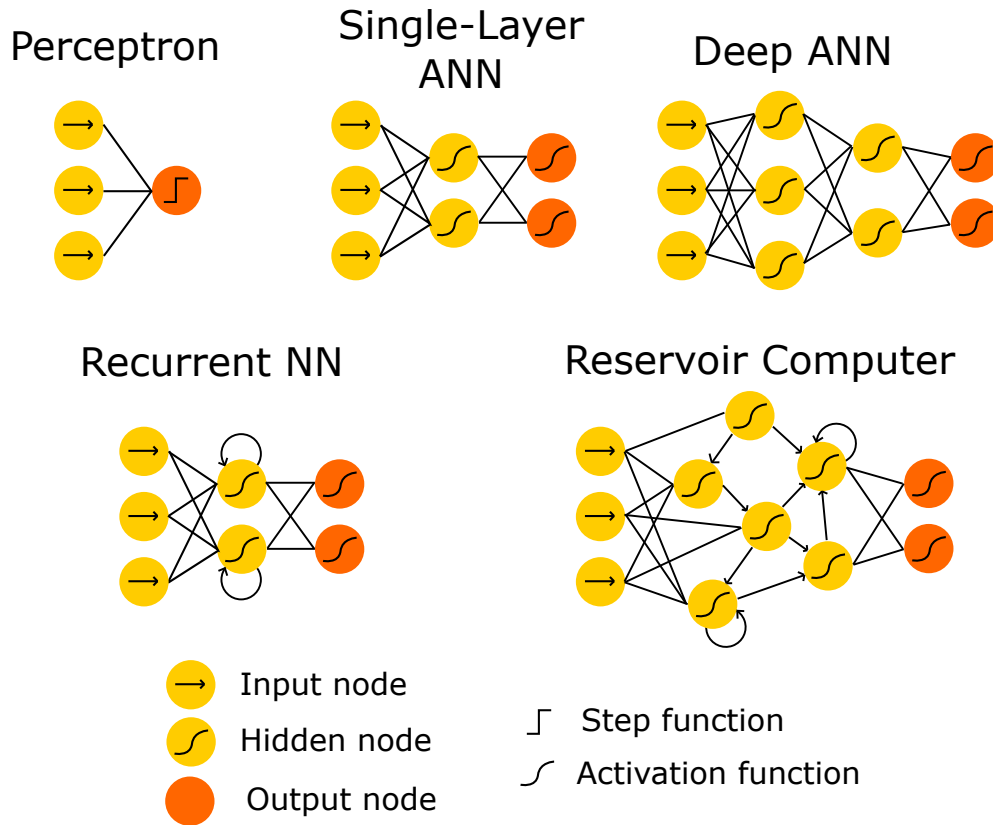


Figure 1.1.3: Different neural network architectures. Nodes represent the nonlinear activation functions applied to the data flowing through the network, while connections represent the weighting, or linear transforms, in between the layers of nodes. The perceptron is the simplest neural network, being used for one-class classification, while the reservoir computer is an architecture that has been commonly used in photonics, having internal weights that are set randomly instead of being trained.

1.2 Neuromorphic Processing

An Alternate Paradigm of Computation

Since the invention of the integrated logic gate, the technology used to build the chips for digital computers has progressed at such a rate that in the 1960's, the CEO of Intel Gordon Moore, made the observation that the number of components included on a single integrated circuit roughly doubled each year. This 'law' meant that the complexity - and thus computational power - of digital electronic integrated circuits rapidly increased, allowing the widespread use of computers in all walks of life. This growth however is limited by the size of integrated transistors and the resulting challenges

such as leakage current and excess heat, and circuits based on digital electronics are reaching a point where development by scaling of component count or clock frequency is becoming more challenging. This offers an opportunity for new computing platforms that do not derive their advantage from simple scaling of integrated chips. Additionally, the implementation of modern AI models - in datacentres and for LLMs - is massively driving up the demand for computational resources, including the usual Graphics Processing Units (GPUs), memory and hard drives, but also electrical power and water (used mostly to remove waste heat from vast banks of processors), with reports of some datacentres potentially using more water than entire cities[17], further encouraging the development of hardware with low power requirements to replace digital computers in AI implementations.

Neuromorphic computing is the name given to a computing paradigm loosely based on, and inspired by, the processes in the brain and nervous system that allows living beings to solve a wide range of problems that we might want computers to do, for example pattern recognition (in vision, hearing etc.) or fine motor control. The brain is composed of millions of individual neuron cells, each of which operate asynchronously and individually, each responding to the electro-chemical inputs passed to them from other neurons, or external sources such as senses, and has very little in common with the typical Von-Neumann architecture of modern digital computers: therefore, neuromorphic engineering seeks to make use of the principles of the operation of the brain to build alternate-architecture computing, processing and sensing platforms. The design of these neuromorphic systems can be done either in the fashion of accelerating existing computing systems - i.e. reducing the load on traditional electronic computing hardware by moving some processes to neuromorphic platforms, or by designing ad-hoc hardware solutions from the bottom up.

Research into neuromorphic platforms has largely focused on electronic hardware, using chips implementing spiking neural networks (SNNs) on large arrays of processors. Notable systems include the SpiNNaker systems (from the University of Manchester), featuring millions of digital processor cores simulating billions of neurons and synapses[18, 19]; BrainScaleS (from Heidelberg University)[20], which uses several hun-

dred analogue-valued neurons, with digitally controlled synapses; IBM TrueNorth[21] and the following NorthPole[22], both of which run over a million simulated neurons, while consuming a relatively low power (several milliwatt); and Intel’s Loihi chips[23, 24] that offers similar specifications, which have been further packaged into the Hala point system, containing over 1000 Loihi 2 chips.

Neuromorphics is also well poised for use in the fields of edge processing and sensing. Edge computing refers to a model of distributed computing power in which data storage and processing remain near to the source of the data and the location where the processed result is desired, in contrast with cloud computing, which involves sending data from sensors to far-off computing centres to make use of powerful processing for portable systems. Neuromorphic platforms provide many of the desirable properties for edge processors: namely very low power consumption, low latency and importantly the integrated implementation of AI solutions[25].

It should be noted that neuromorphic engineering does not purport to seek understanding of or recreate the full range of behaviour and dynamics of biological nervous systems, but instead aims to use the same principles that make neurons and the wider brain system capable of complex computation while remaining incredibly energy efficient.

Networks of Spiking Neurons

While artificial neural networks use real valued signals, and operate without a concept of time across the network (i.e. all weights and activations are calculated sequentially without any explicit timescale), biological networks consisting of neurons (e.g. in our nervous and sensory systems) communicate using sequences of ”action potentials” - discrete, large perturbations in the potential of the neuron’s membrane. Collections of these neurons, each connected to other neurons and able to send and receive signals from their connected neighbours, form systems known as Spiking Neural Networks (SNNs).

SNNs represent and transmit information through discrete spikes (mimicking the action potentials of biological neurons), which can be interpreted either by the firing rate

(rate coding) or by the precise timing of individual spikes (temporal coding) produced by the neurons in the network[26, 27]. This mode of communication and processing offers several advantages: Sparse spike-based activity consumes energy only when a neuron emits a spike, enabling greater computational efficiency; Since information is carried in spike timing rather than in continuously varying signals, it is preserved even in networks where the dispersion of signals (e.g. electrical or optical signals carrying the data) would otherwise reduce transmission capability; Spikes can be regenerated by intermediate nodes, helping to limit noise accumulation and compensate for lossy connections. Additionally, SNNs show the same universal-approximation capability as traditional ANNs, however often requiring fewer nodes in the network[26].

Training these spiking neural networks is a rather more challenging task than for typical ANNs, since the desired output of the SNN is not always well defined, and the transfer function between neurons (i.e. the mechanism that determines when spikes are fired) is not necessarily differentiable, prohibiting the use of well-proven methods such as backpropagation. Several alternate methods have been proposed instead, being inspired by the processes observed in biological neurons as they learn, or constructed specifically to use a learning rule given a set of spike times.

Extensions of the typical backpropagation training methods to SNNs include backpropagation through time (BPTT)[28], a method originally used for recurrent ANNs that implements traditional backpropagation on an 'unfolded' recurrent network; Spike Pattern Association[29], and Spike Layer Error Reassignment (SLAYER)[30], each of which adjust weights according to a post-hoc calculated error term; and surrogate methods that simply use an ANN with differentiable activation functions as a 'stand-in', and transfer the trained weights for use in the SNN. These training algorithms however do not explicitly utilise the temporal dynamics of SNNs, require similar resources as needed for traditional ANNs, and do not necessarily represent the best performance.

Biologically inspired learning methods are mostly derived from Hebb's rule, the observation that neuron cells that fire spikes at similar times tend to become more closely linked to each other[31]. This is often implemented in SNNs using Spike-Timing-Dependent-Plasticity (STDP): adjusting the weights of neuronal synapses according to

Local Training: Parameters of the network should be affected only by information immediately relevant to each parameter, and should not be affected by processes occurring in distant parts of the network.	Non-local Training: All parameters can be adjusted according to information from anywhere in the network. Contrary to local training, this requires an external system to observe the entire network and decide on updates.
Online Training: Training of parameters is done simultaneously to the processing of information, as new data is fed into the network, the weights are updated to reflect best performance	Offline Training: Data is processed, then the adjusting of network parameters is calculated afterwards based on the final outcome, compared to the desired outcome
Hardware Constraints: Some functions cannot be implemented using physical hardware (e.g. when using the intensity of light for computing, negative numbers cannot be used)	Processing Requirements: The hardware components must be able to implement the relevant operations on the input data (e.g. matrix multiplications, nonlinear transforms)

Table 1.1: Features to take into account when designing hardware-based spiking neural networks.

the differences in times of spikes from the neurons before the synapse (the presynaptic, or input neuron) and spikes from the receiving neuron (postsynaptic, or output). Figure 1.2.1 shows one example of the possible ways to adjust weights depending on spike timing. In this example, when the postsynaptic spike closely follows the presynaptic spike, the weight is increased, such that the postsynaptic neuron becomes strongly dependent on the presynaptic neuron, while a presynaptic spike following a postsynaptic spike weakens the link between the two neurons. Further training methods created specifically for SNNs are already in development, making use of the phase of oscillatory neurons[32], the sub-threshold dynamics caused by spikes at synapses[33], or targeting the recreation of particular spiking patterns[34].

Design of hardware-based spiking neural networks has a very wide set of parameters to choose from, including balancing the exact requirements of the networks capability, and the limitations set by the hardware used. Table 1.1 lists a few major points of consideration.

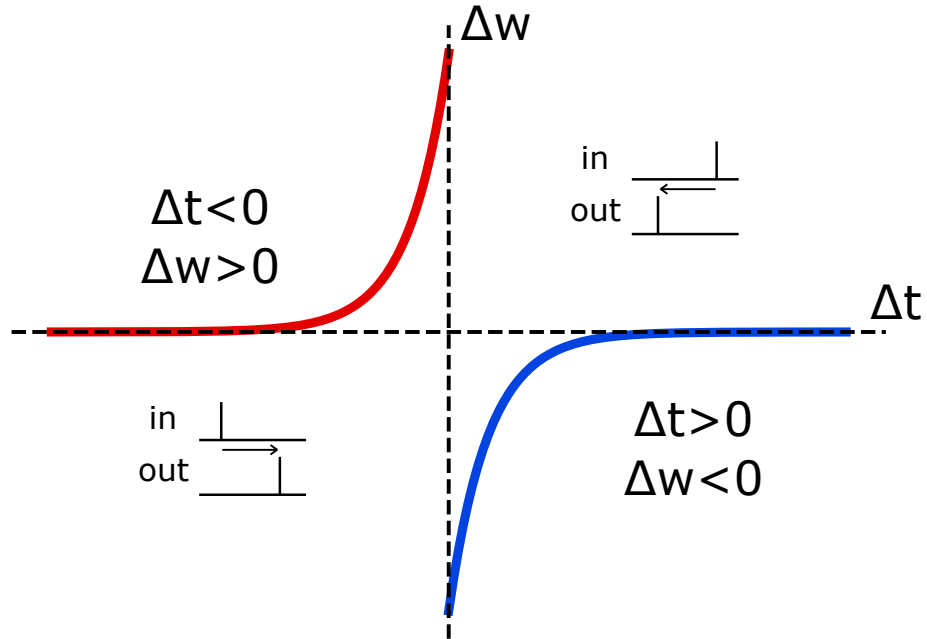


Figure 1.2.1: Adjusting neuron weights using spike-timing-dependent-plasticity (STDP). When the input (presynaptic) neuron spikes before the output (postsynaptic) neuron, the weight is increased, and so the dependence is reinforced. Conversely, if the input neuron spikes after the output neuron, the weight is decreased.

The Dynamics of Excitable and Spiking Systems

A number of models exist for use as the spiking neurons in SNNs, ranging in complexity from mathematical abstractions of neural activity, to detailed dynamical systems.

Excitable and spiking systems form the foundation of information processing in biological and artificial spiking neural networks. Their hallmark is the ability to generate rapid, transient responses—action potentials—when perturbed beyond a threshold, enabling reliable communication across cells and circuits. Understanding the dynamical principles that govern these responses is essential for interpreting neuronal behaviour, constructing accurate computational models, and uncovering the mechanisms by which networks encode and transform signals.

Modelling of biological neurons begins by considering the potential difference across the membrane of a neuron, caused by the concentrations of different ions inside and

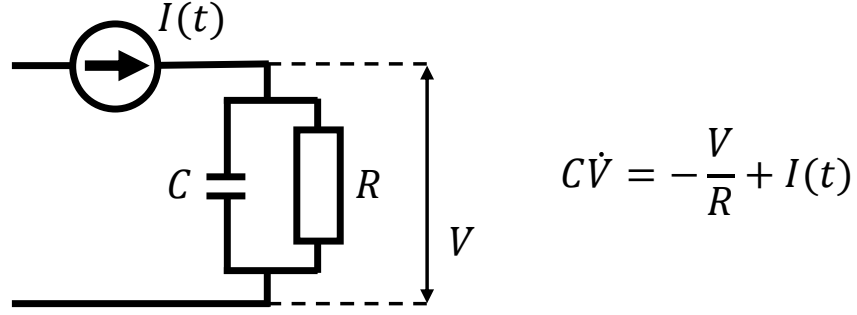


Figure 1.2.2: The effective leaky-integrate-and-fire neuron circuit. Inputs to the neuron are represented by the current source term, that drives charge onto the capacitor, increasing the voltage. When the current is low, the capacitor is slowly discharged via the "leaky" resistive channel. In the artificial spiking neuron model, the neuron is considered to spike when the capacitor voltage reaches a threshold level.

out of the cell. Movements of these ions, through chemical processes or via activity due to neurotransmitters then cause changes in this potential, which can lead to a spike (or "action potential"). The Hodgkin-Huxley model[35] was introduced based on the measurements of these ionic currents in the neurons of a giant squid. In this model, four differential equations are used to model the voltage and current across the membrane, which together display a range of behaviours seen in biological neurons. Further models, such as the FitzHugh-Nagumo (FhN) model[36] reduce the complexity of the Hodgkin-Huxley model by combining several current terms and making simplifying assumptions on the different timescales.

The typical model used is known as the "leaky-integrate-and-fire" (LIF) neuron. This emulates the neuron's membrane as a parallel capacitor-resistor circuit, with the input signal represented by a current source. The input current accumulates charge on the capacitor, increasing the voltage ("integrating" the input), while the parallel resistive term allows the voltage to decay to zero over time (the "leak" current). When some threshold voltage is reached, the neuron is considered to emit a spike, and the voltage is reset to zero, and held constant for a short time, in which the neuron does not respond to new inputs (the "refractory period").

Further neuron models often extend this, one example being the Resonate-and-Fire (RaF) neuron. In this case, instead of integrating the input signal, the neuron responds

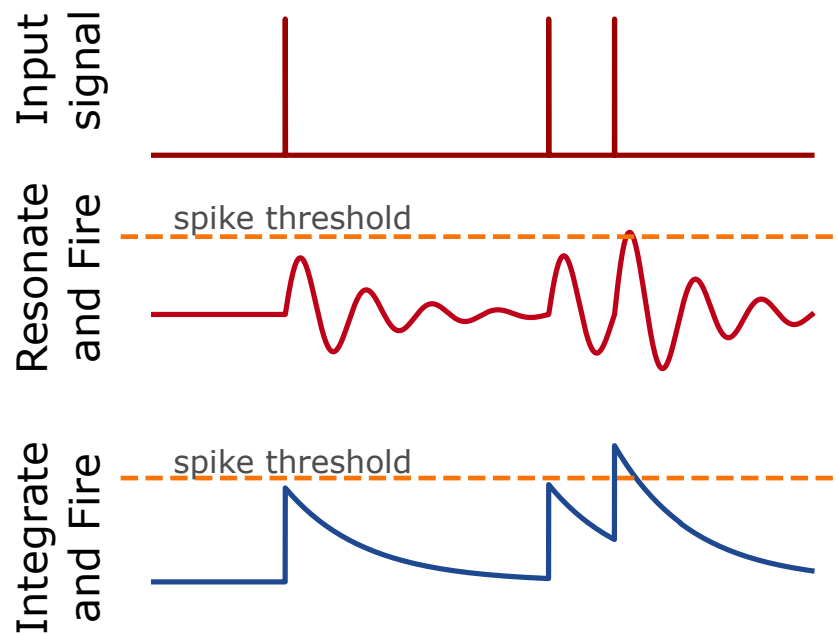


Figure 1.2.3: Comparison of the response in Resonate-and-Fire and Leaky-Integrate-and-Fire neurons to a given input. When the responses to the signal sum constructively the result can cross the threshold required to spike.

in an oscillating fashion with a set frequency and decay rate. Input signals with strong frequency components near to the resonant frequency of the neuron will lead to a stronger response and therefore more spiking activity. Alternatively, when considering neurons connected to each other, the RaF neurons will fire preferentially to input spikes that are separated by a specific time interval, rather than being sufficiently near to each other, as for the LIF neuron. The responses of LIF and RaF neurons to input spikes are depicted in figure 1.2.3, showing the sub-threshold response to an input, and how separate input events can combine to produce an output spike.

While the abstracted LIF and RaF can faithfully represent the interactions between neurons and their inputs, the outputs are only representations of the spike trains (i.e. spikes are described by the time at which the neuron’s potential exceeds its threshold), and the refractory period of the neurons is ‘enforced’ externally. Dynamical models (such as the Hodgkin-Huxley or FitzHugh-Nagumo models) instead implement the features of neurons in the intrinsic dynamics themselves. These dynamical systems

feature phenomena known as bifurcations, where a small change in a parameter in the model (e.g. driving voltage) can cause a large-scale change in the global dynamics, including the creation and destruction, or change in stability of fixed points, and the emergence of new (possibly periodic) trajectories. Excitability refers to a particular set of behaviours where a temporary perturbation to some system parameter can cause the system to travel along a defined and repeatable trajectory before returning to its original state, that depends only on the input perturbation exceeding some threshold (be it in amplitude or duration).

Biological neurons can be grouped into classes based on the character of their spiking responses. Class I neurons can generate constant spiking responses over a wide range of firing frequencies, with the frequency determined by the strength of the input stimulation. This behaviour is generated by a saddle-node on invariant circle (SNIC) bifurcation and is commonly observed in cortical excitatory pyramidal cells.

In contrast, Class II neurons produce trains of spikes within a narrower, preferred frequency band, and their firing rate is less sensitive to changes in input perturbation strength. They are typically modelled using systems that exhibit a Hopf bifurcation, in which a stable fixed point is destabilised and the system trajectories approach a limit cycle, demonstrated in figure 1.2.4. A spike is fired by the system completing a period of the limit cycle before returning to the original fixed point, which may have re-stabilised if the input perturbation is removed, or may still be unstable in which case a second spike (orbit of the limit cycle) is fired.

1.3 Photonics as a Platform for Neuromorphic Processing

Photonics refers to the use of light, in the form of photons for tasks such as communications, sensing and computing. The most prominent example of photonic technologies in everyday life is the use of optical fibres for carrying communications traffic - where previously these were carried over electric telegraph cables. The use of photons over electrons is primarily motivated by the potential for much higher modulation bandwidths, in addition to the lower crosstalk and lower sensitivity to external interference

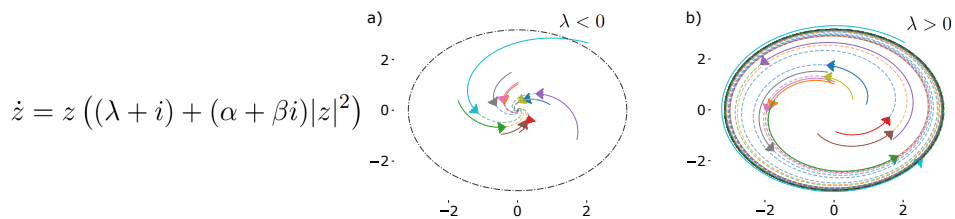


Figure 1.2.4: Two behaviours of a system exhibiting a Hopf bifurcation. Trajectories are plotted for a randomly selected set of starting points. The sign of the parameter λ determines the global behaviour of the system: when $\lambda < 0$ (a), all trajectories spiral in to the stable fixed point at $(0,0)$, while for $\lambda > 0$ (b) the fixed point is destabilised and the trajectories exponentially approach a stable limit cycle (plotted as a dashed grey line). The point at which $\lambda = 0$, where the global dynamics change, is referred to as the bifurcation point.

that is possible over optical fibres than cables.

Research into the use of photonic devices for applications to neuromorphic processing has been growing rapidly over the past several years, covering a wide range of devices and different methods of use[37]. Typically, these approaches do not rely entirely on photonic platforms, but use a combination of opto-electronic systems in which some operations are handled by conventional electronic systems, and others by photonics, in such a way to glean an advantage in domains such as processing speed, bandwidth or power efficiency.

In spite of the relative infancy of photonics with respect to electronic technologies, when it comes to neuromorphics there are several outstanding benefits to the use of light:

- High bandwidth: Optical systems can operate at bandwidths far above those attainable with electronics[38] (typically upwards of several Gigahertz or even Terahertz). The parasitic capacitance and inductance present in all electronic systems limit their bandwidth (i.e. very short pulses cannot be propagated quickly or efficiently) and the speed at which signals can be transmitted.
- Low intrinsic communication losses: transmitted light - through optical fibre or free space - loses very little power. This leads to a reduced power requirement, and lower heating due to waste power, allowing for more dense integration of

optical components.

- Passive processing: The dynamics of light, and the interactions between light and matter can be used to implement functions required for computing, without the need to explicitly drive any particular process. For example, the free-space diffraction of light can be used to calculate the Fourier transform, by modulating the starting intensity profile of the light with the function to be transformed.
- Wavelength Division Multiplexing (WDM): A single optical fibre can support a range of wavelengths of light, with little to no crosstalk between each wavelength mode, allowing for very wide broadband multiplexing, in which each wavelength carries different information.

There are however, a few difficulties in the implementation of integrated photonics platforms for computing:

- Non-interacting signals: While the minimal interaction of light, with matter or with other light, is beneficial for reducing power losses during transmission, it makes implementing components such as logic gates difficult.
- Manufacturing: Integrated Photonic Circuits are an emerging technology[39], which can make use of existing semiconductor foundries, however there still remains challenges regarding integration of a full complement of photonic components. These photonic components can also be limited in size by the wavelength of the light used, making the chip density lower than that of electronics.
- Optical-Electrical Conversion: The conversion of signals from electrical to optical or vice versa is energy inefficient, and can put limits on bandwidth or modulation speed.

Photonic processing systems can be used to replace sections of traditional electronic based computers e.g. as a photonic accelerator[40], giving a particular benefit such as an increase in processing speed to an existing computer, or as self-contained processing systems of their own, with electronics used only to interface between digital

computers (with an operating system etc.) and the photonic machine. Here I will list some previous works that have made use of different photonic technologies for AI and neuromorphic platforms.

Lasers

Lasers are the central component of most photonic platforms - being the way to produce coherent light that can be carefully controlled. The light emitted by a laser is strongly collimated, maintains a stable phase relationship through space and time, and exhibits a strongly controlled spectrum, making it ideal for manipulation and application for optical computing. A laser requires two essential components: an active or gain medium and an optical resonator that feeds light back into the gain medium using reflective elements. The fundamental mechanism enabling lasing is stimulated emission, in which an incoming photon triggers the emission of an additional photon that is identical in phase, direction, and energy, producing a coherent copy. By building a population inversion of charge carriers in the energy levels of the gain medium, the probability of stimulated emission is increased. As photons are circulated repeatedly through the medium as standing waves within the resonator, they undergo amplification on each pass. When the net photon gain exceeds losses in the system (from absorption or the useful emission of light out the laser), sustained laser emission is achieved.

The neuromorphic applications of lasers range from use simply as sources of coherent light, to be modulated and facilitate optically downstream processes; to the active elements themselves for the neuromorphic platform. Particularly of interest for the development of integrated photonic platforms for neuromorphics are semiconductor lasers (SLs)[41]. A range of different lasers have found use as neuromorphic platforms, including ring lasers [42], microdisk lasers [43], and several species of Vertical-Cavity Surface-Emitting Laser (VCSEL)[44]. VCSELs exhibit rich nonlinear dynamics, particularly under optical injection, which have found use in photonic reservoir computers[45, 46] due to the simple hardware requirement and high modulation speeds (up to several GHz). Novel forms of these lasers, such as Large-Area (LA) and spin VCSELs have found similar use[47–49]. LA-VCSELs feature a wide cavity that can support multi-

modal emission and complex dynamics and interactions between the cavity light and the distribution of carriers over the surface of the laser, while spin-VCSELs[50] have strong interactions between the different polarisations in the laser cavity and the spin of the carriers.

Microring Resonators

Microring resonators (MRRs) are integrated components formed of micron-scale rings that support optical whispering-gallery modes, evanescently coupled to nearby optical fibres. Light passing down the optical fibres leaks into the ring, with light with the same wavelength as the resonant mode of the ring coupling very strongly. This resonant wavelength can further be tuned by thermal and electric control - using any method of perturbing the refractive index of the ring material.

The wavelength selectivity of MRRs makes them ideal for computing on WDM signals, since the ring acts as a 'weight', scaling the power that passes by a factor dependent on the wavelength. MRRs can also be used as space-multiplexed weights for neural networks[51, 52], adjusting the amplitude of signals passing through them, or even as the nonlinear elements themselves in an ANN or reservoir computer[53, 54]

Optical spiking has also been demonstrated in MRR systems[55, 56]: Enhanced two photon absorption inside the ring resonator leads to the release of free carriers and heat into the ring, which can cause a change in the resonant wavelength, possibly leading to a change in the power coupled into the ring, setting up a feedback system on the timescale of these effects - upwards of 10s of nanoseconds. This feedback loop, in conjunction with adequate input light can lead to complex self-pulsing and deterministic, triggerable optical spiking. A recent work has also demonstrated a large-scale spiking neural network platform that leverages the functionality of MRRs for high energy efficiency and low hardware complexity[57].

Optical Modulators

Light modulators are optical components that can be used to programmatically change the properties of light (e.g. phase, amplitude) passing through them, based on some

electronically-controllable feature. For example, the pockels effect present in materials with a strong second-order electric susceptibility allows for varying the refractive index of a slab that material by applying a voltage across the slab, which in turn can be used to introduce a controllable phase shift to light passing through. In a Mach-Zehnder modulator (MZM), this phase shift is used to cause interference at the output port of an interferometer, allowing the modulation of the amplitude of the light in time.

Spatial Light Modulators (SLMs) can be built using liquid crystal pixels, that allow modulation of the intensity profile of light passing through in a similar manner to optical displays. Digital Micromirror Displays, consisting of large arrays of electronically-controllable micromechanical reflectors, can also control the spatial distribution of light. SLMs and DMD modulators often operate with modulation speed up to several KHz, and due to the very high parallelism (e.g. millions of independent pixels) can represent very high processing-power components.

Typically, modulators are used in photonic neural networks to implement the weights of the linear transforms between the layers of nonlinear nodes (or as the active element themselves in e.g. an Ising machine[58]), due to the fine, individual tunability of the pixels. In [59], an SLM was used to apply weights for a spiking neural network built using an LED and camera, being able to emulate 40000 spiking neurons with simple hardware feedback, while in [60], a DMD array was used as the output layer weights for a reservoir computer based on a wide-area VCSEL, with training done by flipping the pixels of the mirror array one-by-one until strong performance is found. Arrays of MZMs have been applied similarly [61, 62], allowing low power cost matrix multiplications, and even an in-hardware backpropagation training system [63]. Further cases have made novel use of modulators as active elements in all-optical or opto-electronic feedback systems: [64] uses the sinusoidal transfer function of an MZM in a feedback loop (feeding the photodetector voltage generated by the modulated light into the control port of the modulator) as analogue spins for use in an Ising machine, able to implement several hundred spins through time multiplexing; in [65], several MZMs and phase modulators are connected (along with laser sources and amplifiers) to form a deeply connected photonic reservoir computer running on frequency multiplexed signals.

1.4 Thesis Outline

This thesis is structured into 6 Chapters, with two of them focussing on neuromorphics using Vertical-Cavity Surface-Emitting Lasers (VCSELs), and the other two on Resonant Tunnelling Diodes (RTDs), and a final summary and outlook.

Chapter 2 will describe the structure and function of VCSELs, and the experimental apparatus used to drive and study the lasers for the work included in chapters 2 and 3. The dynamics of VCSELs that lend them to being used as photonic spiking neurons will also be explored, demonstrating typical neuronal behaviours such as leaky-integrate-and-fire spiking. Following that, Chapter 3 will make use of these spiking VCSEL neurons to construct photonic spiking neural networks (SNNs), capable of complex nonlinear tasks including classification and time-series prediction. The algorithms supporting the hardware implementation will be discussed, and novel training algorithms for use with the photonic spiking neural networks will be introduced.

The use of Resonant Tunnelling Diodes (RTDs) as opto-electronic spiking neurons will be explored in chapters 4 & 5. Chapter 4 introduces the device and theoretical model of the RTD neuron circuit, explores various dynamical behaviours of this model and dives into how this can be used as an artificial photonic-electronic spiking neuron. The typical features such as integrate-and-fire, thresholding and refractoriness are shown and characterised using numerical simulation. Finally, the experimental setup used to investigate RTD-based photonic-electronic neurons is shown, along with experimental results of spiking RTD neurons. Chapter 5 then uses spiking RTD neurons (in simulation and experimentally) for several neuromorphic systems, including a spatially-multiplexed Extreme Learning Machine (ELM), recurrent spiking memory structures, and an event-based time-series edge detector using dual electronic and optical modulation.

Chapter 2

Spiking VCSEL-Based Photonic Neurons

This chapter will introduce the class of semiconductor lasers known as Vertical-Cavity Surface-Emitting Lasers - from here referred to as VCSELs, and explore how they can be used as photonic spiking neurons[44, 66–71].

VCSELs are distinguished by the emission of light perpendicular to the surface of the wafer on which they are grown, rather than from the edge, as in conventional laser diodes. Since their development in the late 20th century[72], VCSELs have become ubiquitous components in optical communications, sensing and emerging photonic computing technologies[73]. The vertical structure and thin cavity[74] allow for low threshold currents, high modulation bandwidths, and reduced manufacturing costs due to the capability for on-wafer testing[75–77]. Additionally, the possibility of easy integration in wider two-dimensional array structures makes them particularly attractive for scalable optical systems[78]. Previous research has revealed that VCSELs can exhibit rich nonlinear dynamics, such as periodic oscillations and chaos[79, 80], and crucially for neuromorphics, polarisation switching and excitability[81], that mimic the behaviour of biological neurons.

2.1 Vertical Cavity Surface Emitting Lasers

VCSELs are fabricated through molecular beam epitaxy or metal-organic chemical vapour deposition, growing layer by layer the distributed Bragg reflectors (DBRs) that form the vertical mirrors for the lasing cavity, the active layer between them that amplifies the laser light, and electrodes on the top and bottom of the stack that allow for electrical pumping[82]. This vertical stack structure, shown diagrammatically in figure 2.1.1, brings a few advantages compared to edge-emitting lasers, most notably that testing can be done during manufacturing, allowing for errors to be corrected, reducing material and time waste and so the final production cost. The beam divergence is typically lower[82], and the thin active region makes single longitudinal-mode emission easy to achieve, as well as allowing high-frequency modulation for optical communications[77]. VCSELs have thus found themselves used in a range of applications: fibre optic communications (due to their high modulation bandwidth and efficiency), sensing devices (e.g. spectrometers, lidar), optical clocks, and consumer devices such as computer mice or structured light scanners used for face recognition and 3D scanning[83–86].

The fundamental emission mode of the VCSEL cavity is usually a TEM mode polarised along one of two orthogonal directions (determined by the exact shape of the cavity), with a second, typically much weaker emission mode polarised in the other direction. These two modes show slight differences in their resonant wavelengths (approx. 100 pm) due to birefringence in the cavity material[50, 87]. The interactions between the polarisation modes and electronic states can lead to rich non-linear interactions with driving current and injected light.

VCSELs based on GaAs were the first to be built, operating at a wavelength of 850 nm [72, 82]. More recently, longer wavelength VCSELs have also been developed, based on indium alloys, allowing lasing in the telecom "O" (1300 nm) and "C" (1550 nm) bands[88], ideal for long range, high speed communication.

The VCSELs featured in this work are monolithic devices with an InGaAs quantum well based active layer and InAlGaAs/InAlAs Bragg reflectors. These were sourced

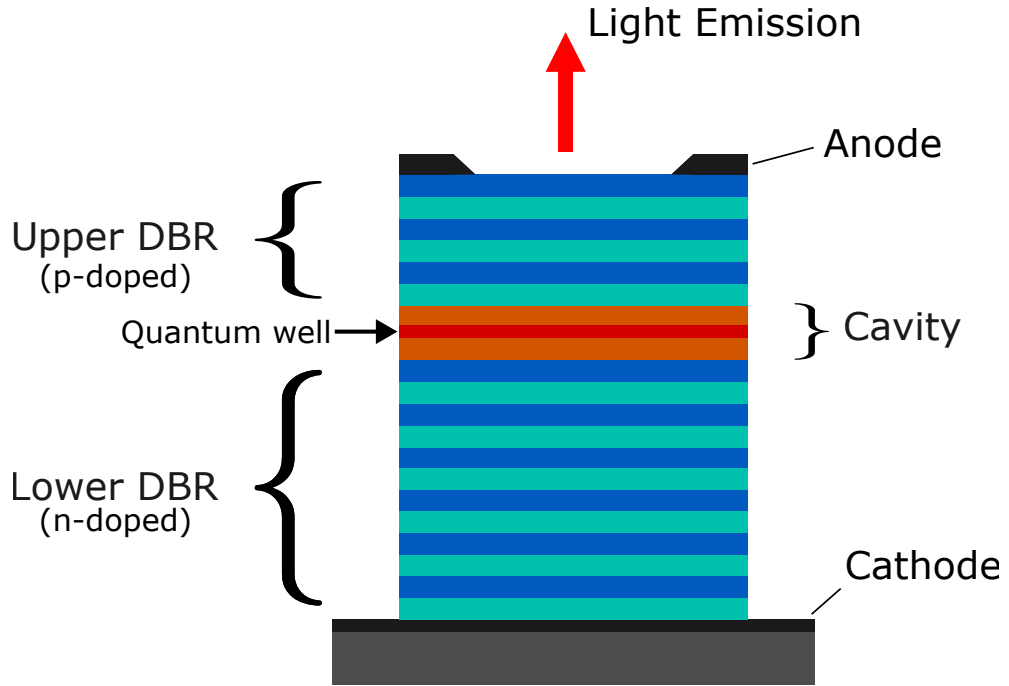


Figure 2.1.1: Schematic of the typical structure of a VCSEL. Each part of the laser (two surrounding DBR mirrors, and the cavity) are grown vertically one at a time. The DBRs are formed of InAlGaAs/InAlAs, with the top and bottom DBRs doped to enable current flow into the cavity. In the cavity are InGaAs quantum dots, that act as the active medium amplifying the light.

commercially from the provider RayCan[89].

The spectrum of the VCSELs used in this work typically shows two peaks - as a result of the two birefringence split lasing modes, however, the subsidiary, lower power peak can often show as a 'shoulder' on the stronger mode rather than a distinct peak. Figure 2.1.2 shows the spectra of two VCSELs operating at (a) 1300nm and (b) 1550nm, with the dominant and subsidiary modes marked. Notably, the subsidiary mode of the 1550 nm VCSEL is at a lower wavelength than the dominant mode, while the opposite is true of the 1300 nm VCSEL.

2.1.1 Optical Injection Locking of VCSELs

When two similar oscillators (e.g. mass-spring systems, pendulums, or laser cavities) are allowed to influence each other, energy can be exchanged between the two, affecting the phase and frequency dynamics of each oscillator. If the frequencies of each oscillator

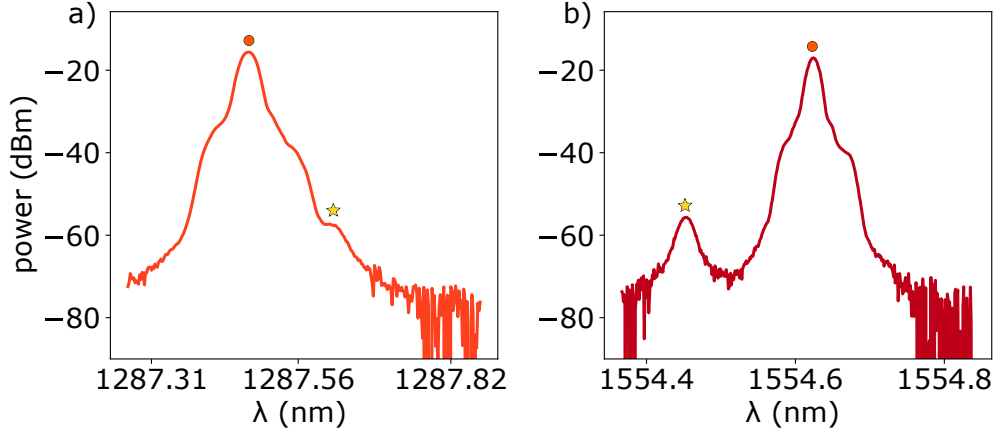


Figure 2.1.2: Typical spectra of VCSELs at 1300 nm (a) and 1550 nm (b), each driven with a bias current of 3.5 mA. The red circles mark the dominant emission modes of the VCSEL, while the yellow stars mark the secondary (subsidiary) modes, that are targeted for optical injection induced spiking dynamics.

are near enough to each other, it is possible that they both exactly synchronise and oscillate at a common frequency, a mechanism known as injection locking[90].

In semiconductor lasers such as VCSELs[91], the exchange of energy that leads to injection locking occurs due to the α enhancement factor of the laser cavity. The α factor of a semiconductor laser describes the enhanced linewidth broadening due to the refractive index of the cavity material being dependent on the density of charge carriers, which in turn can vary due to pumping and the emission of light[92]. This coupling links the amplitude of light in the cavity to the phase, leading to the additional widening of the emitted spectrum. This effect is utilised for injection locking by driving light from an external laser into the VCSEL's cavity, such that the amplitude of the injected light is able to modulate the phase, and thus the dominant frequency, of the VCSEL's emitted light.

Early theoretical[93] and experimental[94] studies were made into the dynamics of injection locked VCSELs under varying wavelength detuning and injection power, and these were able to show regions of stable fixed states, oscillations, and chaos. Injection locking occurs in regions of the parameter space near to these stable states of the VCSEL[95], and at the boundaries of these regions, bifurcations can occur that lead to

rich nonlinear dynamical behaviour.

2.2 Optical Spiking VCSEL Neurons

This section will demonstrate the use of VCSELs under optical injection as artificial optical spiking neurons, and showcase the various dynamics that allow these lasers to be used in photonic neuromorphic computing systems.

Light-triggered excitable dynamics exist in VCSELs at the boundary of injection locking - the ranges of injected power where more power leads to guaranteed locking of the VCSEL, while decreased injection power can lead to unlocking and free dynamics. At this boundary there are a variety of bifurcations between different stable states and limit cycles [96, 97], leading to a range of nonlinear dynamics, such as polarisation switching. Neuron-like optical spiking in VCSELs has been demonstrated a number of times over the past ten years by our group based on polarisation switching dynamics of injection locked VCSELs[81, 87, 98], operating at 1300 nm and 1550 nm . As the VCSEL system passes over this bifurcation point, the transition from one stable state (locked emission) to another (unlocked emission) leads the VCSEL state on a trajectory through the state space that involves a 2π shift in the phase of the emitted light, and a short ($\sim 100\text{ ps}$) rapid increase in the emission power, that manifests as a neural-like optical spike.

2.2.1 Experimental Setup

Optical spiking VCSEL neurons were investigated in this work using the experimental setup described in Figure 2.2.1: A tunable laser operating in the same wavelength range as the VCSEL is used to provide light with the right wavelength and power for optical injection locking of the VCSEL, i.e. several μW of power, with a frequency slightly negatively detuned with respect to one of the VCSEL's modes; the amplitude of the light reaching the VCSEL is intensity modulated in time using a Mach-Zehnder modulator (in combination with a variable optical attenuator); The light is then injected into the VCSEL, and the subsequent output light from the VCSEL is sent to an optical spectrum

analyser and a photodetector for analysis. Additional components such as polarisation controllers are used to match the polarisation of light in the optical fibres to the modes of the VCSEL, and to couple with low optical losses to the MZM. An optical isolator was also used at the output of the tunable laser to avoid damage from reflections entering the TL's cavity. A photograph of the setup described by the schematic is shown in figure 2.2.2.

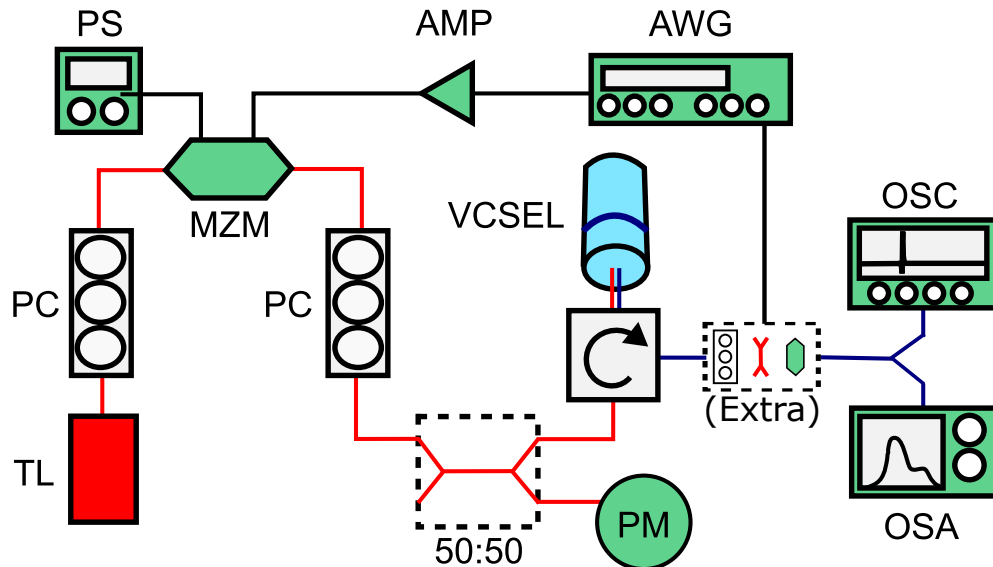


Figure 2.2.1: Diagram of the experimental setup used for investigating optical spiking in VCSEL neurons. Light from the tunable laser (TL) is amplitude modulated by the Mach-Zehnder modulator (MZM) and has its polarisation controlled by the polarisation controllers (PC), before being injected into the VCSEL via an optical circulator. The output of the VCSEL is then fed to an optical spectrum analyser (OSA) or photodetector and oscilloscope (OSC). Voltage signals generated by the arbitrary wave generator (AWG) are used to control the modulators, along with a DC bias from a power supply (PS). The average power of the injected light can also be measured using a 50:50 fibre splitter leading to a power meter (PM).

The VCSELs used were commercially sourced telecom wavelength (1310 nm and 1550 nm) TO-56/TO-90 packaged fibre pigtailed VCSELs, that could be connected to further optical fibres by physical contact or angle physical contact connections. The devices were controlled using Thorlabs LDM90 or LDM21 mounts, that could apply a pumping current to the laser and also control the temperature, allowing stable operation of the VCSELs. The bias current to the mounts was provided by Thorlabs LDC205C

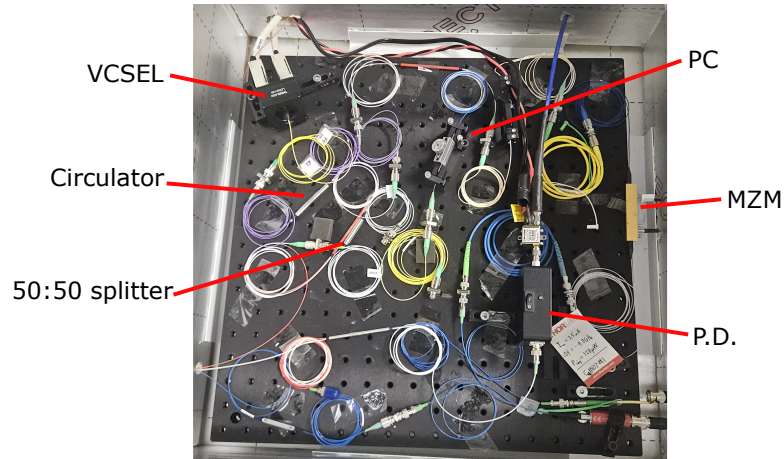


Figure 2.2.2: A photograph of the experimental setup as described by the schematic in figure 2.2.1. The VCSEL can be seen in the top left, on a Thorlabs LDM21 temperature and current control mount.

laser drivers, and the temperature of the mount was set using a Thorlabs TED200C temperature controller. Temperature was kept constant at 293K (room temperature) for all experiments, while the typical bias current used was on the order of $< 10\text{ mA}$, to exceed the typical threshold current of $1 - 2\text{ mA}$ of the VCSELs used in this work.

Two different tuneable laser models were used for this work, one at 1300 nm and one at 1550 nm . The TL at 1300 nm , was the TSL-120 from the manufacturer Santec, which is tuneable in the range $1260\text{ nm} - 1360\text{ nm}$ with a wavelength resolution of $\sim 1\text{ pm}$. At 1550 nm the TL used was the WSL-110 from Santec, with a range of $1527\text{ nm} - 1567\text{ nm}$ at a wavelength resolution of $\sim 0.1\text{ pm}$. Each laser could emit upwards of 10 mW constant wave light, allowing a suitable range of powers for optical injection locking, even when considering the additional losses in the intermediate optical components in the setup.

Mach-Zehnder modulators (MZMs) are interferometers in which one arm contains a Pockels cell, consisting of a material such as Lithium Niobate that shows a strong change in refractive index when there is a voltage across it (the Pockels effect). This allows controlling the phase change of light in the Pockels cell arm by applying a voltage, thus varying the amplitude of the light at the output when the interferometer arms are recombined. Since the effect of a voltage on the Pockels cell is to cause a

phase shift in the light passing through, the transfer function (i.e. output amplitude v.s. input voltage) is sinusoidal. The MZMs used in this work were JDS-Uniphase 10 Gb/s integrated amplitude modulators and Thorlabs LN81S FC 10 Gb/s intensity modulators, which operate with low optical losses in the $1525\text{ nm} - 1605\text{ nm}$ wavelength range (however losses in the 1300 nm range were higher, though not impeding the system). The half-wave voltage (i.e. voltage that causes a π phase change, and so the voltage difference between maximum and minimum transmission) of the modulators was 6 V , allowing for full-range modulation without needing large voltage amplification, and can be modulated at rates up to 10 Gb/s .

The voltage signals used to modulate the MZMs were provided by a Keysight M8190A 12 GSa/s Arbitrary Wave Generator (AWG). This is capable of producing arbitrary waves with amplitude 1 V peak-to-peak with 12 bit precision, allowing modulation of light with pulses on the order of approx. 100 ps . This can be controlled using manufacturer-provided software, or through commands issued by custom-built python scripts. Since the AWG output is lower than the range needed for the MZMs, an RF electrical amplifier (Minicircuits ZX60-14012L-S+) was used. This has 15 dB voltage amplification at frequencies between 300 kHz and 12 GHz , ideal for use with the VCSELs used in this work.

To measure the output of the VCSEL, their emitted light was sent to an optical spectrum analyser (OSA), and to a fast photodetector (P.D.), which was in turn connected to a real time oscilloscope (OSC). The OSA used was a Yokogawa AQ6370D. This was used to measure the peak wavelength of the VCSEL and the tunable laser in turn, giving a measurement of the detuning of the injection light, and could be used to detect optical injection locking, as the distribution of the VCSELs output power between the two modes changes when locked.

The photodetector used was a Thorlabs PDA8GS amplified photodetector. This has a bandwidth of 9.5 GHz and can operate with wavelengths from 950 nm to 1650 nm . The P.D. was connected via SMA cable to a Rohde&Schwartz RTP real-time oscilloscope for recording and analysis. This oscilloscope has a sample rate of 40 GSa/s and bandwidth of 16 GHz (on a single channel), and could be connected to a PC by USB

cable for automated recording.

Other optical fibre components used were polarisation controllers, splitters/couplers and optical circulators. Fibre-optic polarisation controllers consist of coils of optical fibre that from fractional waveplates depending on their rotation. Stress induced birefringence in the optical fibre allows controlling the polarisation change at the output of the controller by twisting the fibre. These are used to match the polarisation of light from the tuneable laser to the MZM, and subsequently to the modes of the VCSEL (since the optical fibres used were not polarisation maintaining), to minimise losses and reduce the amount of optical power needed (from the TL) for injection locking. Fibre-optic splitters (that can also act as couplers) split their input optical power of two input optical fibres between two output fibres at fixed ratio (e.g. 50:50, 70:30 or 99:1). Polarisation dependent optical splitters can separate light into two orthogonal polarisation modes, one into each output. An optical circulator is used to separate the input light to a VCSEL from the light being emitted by the VCSEL. Since the VCSEL is connected to an optical fibre only on one end, the VCSEL's emitted light must travel down the same fibre as the light to be injected. The optical circulator is a three-port component that uses Faraday rotation and birefringence to allow light to travel only from one port to the next (i.e. light input into port 1 is output at port 2). Thus the injection line is connected to port 1, the VCSEL to port 2 and the collected output to port 3.

This experimental setup is used throughout this work for demonstration of spiking in the VCSEL neuron, and in the following chapter for the construction of photonic spiking neural networks (p-SNN).

2.2.2 Excitable Optical Spikes

Figure 2.2.3 shows the deterministic triggering of an optical spike in a VCSEL neuron using a 200 *ps* long drop in the optical power of injected light, in VCSELs with wavelengths near 1300 *nm* and 1550 *nm*.

In the case of the 1300 *nm* VCSEL, the laser was driven with a current of 4.055 *mA* (above the threshold of 1.4 *mA*), and injected with light detuned by -3.2 *GHz* from

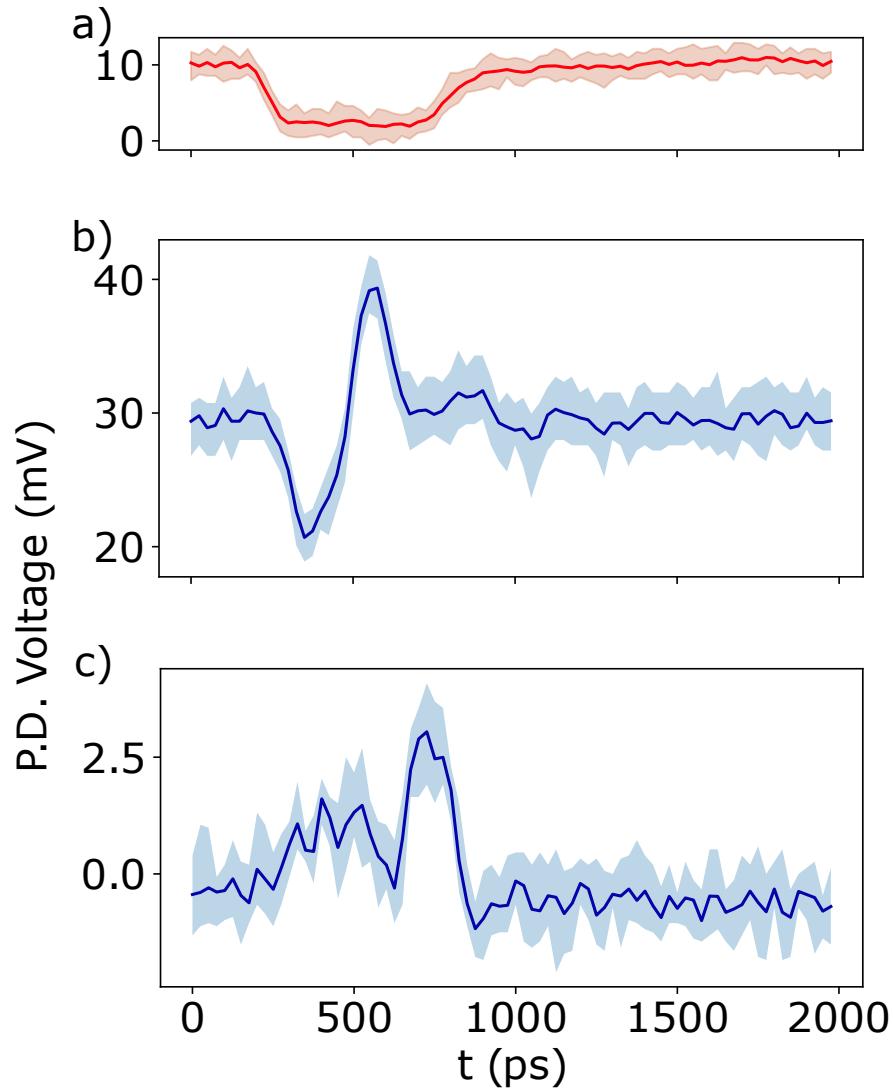


Figure 2.2.3: An excitable neuron-like spike emitted by a VCSEL under modulated injected light. a: The input light used to trigger a spike. The decrease in injected power causes unlocking, which consists of the VCSEL undergoing a limit cycle transition that manifests as an optical power spike. b: The output of a 1300 nm VCSEL. The VCSEL was driven with 4.055 mA and had the peak of its subsidiary mode at 1287.842 nm. 174 μ W of light at 1287.86 nm (a detuning of -3.2 GHz) was injected into the VCSEL and modulated with a 200 ps drop in power. c: The output of a 1500 nm VCSEL. The VCSEL was driven with 4.5 mA, with the subsidiary peak at 1555.625 nm. 41 μ W of 1555.65 nm light was injected (a detuning of -3.1 GHz), and again modulated with a 200 ps drop in power.

the 1287.842 *nm* subsidiary mode. Using an injected light with a power of 174 μW , this locked the VCSEL to the external injection, enabling the excitable spiking response. Similarly, for the 1550 *nm* VCSEL, a driving current of 4.5 *mA* was used (for a threshold of 1.8 *mA*), and injected at -3.1 *GHz* from the 1555.625 *nm* peak, with 41 μW . When the injected light was modulated with a fast 200 *ps* drop of 80% in its amplitude, the VCSEL relaxes away from the injection locked state, and emits the optical spike, seen in the 1300 *nm* VCSEL trace as a decrease followed by increase in power, and in the 1550 *nm* VCSEL as a pure increase in output, across ~ 200 *ps*. Traces of these optical spikes are shown in Figure 2.2.3. These traces show the result averaged over 10 repetitions of the same input, along with the minimum and maximum voltages recorded by the photodetector during the spike. The shape of the spike shows strong consistency across all repetitions, demonstrating the reliable excitable response of the VCSELs.

A rise in the output power of the VCSEL is considered a spike if it follows the particular shape (as seen in Figure 2.2.3), with the peak power exceeding the noise floor of the VCSEL, and if it occurs specifically as a result of a decrease in injected power (rather than by e.g. an increase in driving current) As described, the optical spike is the result of polarisation switching in the VCSEL during the bifurcation that happens on unlocking.

Each polarisation component of the VCSELs output can be observed at the outputs of a polarisation dependent splitter, using a polarisation controller to isolate each component. The results of this split is shown in Figure 2.2.4.

Since one polarisation component undergoes only an increase in the output power during a spike (while the other only decreases), this can be isolated to make detecting a spike easier: the full-width at half maximum (FWHM) of the isolated increasing component is higher (almost double) than the combined spike, potentially allowing for detection of the spike with a slower sampling speed, but also further modulation of the VCSEL output. Two orthogonal polarisation components of the VCSEL's output during a spike are shown in Figure 2.2.4, demonstrating the separate rise and fall in power of each component.

The typical neuronal behaviours of thresholding, integration, and refractoriness that

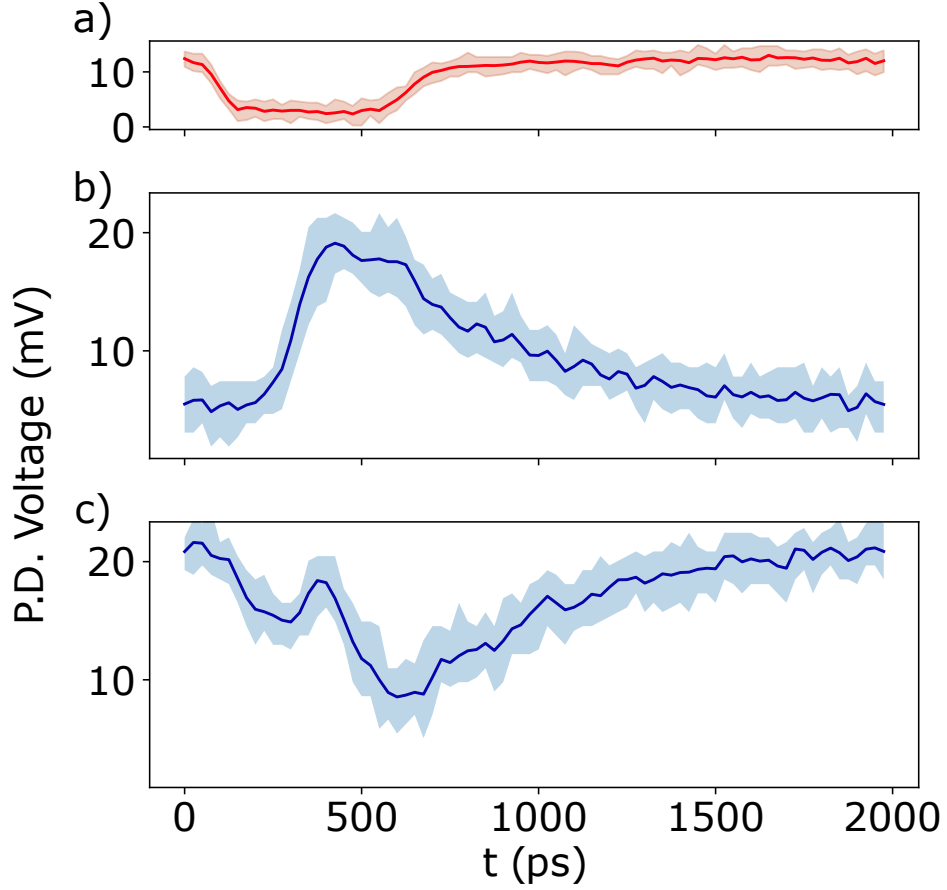


Figure 2.2.4: Orthogonal polarisation components of the output of a 1300 nm VCSEL during a spike. a) The optical input to the VCSEL. The drop in power triggers the spike. b,c) The output power of the VCSEL, split by polarisation to capture the component that increases (b) and decreases (c) during a spike. This separation shows that the dynamics of the spike continue for some time after the initial 'spike' of the polarisation-combined output is observed, which contributes to the refractory period of the VCSEL neuron. The VCSEL was driven with 4.055 mA and had the peak of its subsidiary mode at 1287.842 nm . $174\text{ }\mu\text{W}$ of light at 1287.86 nm (a detuning of -3.2 GHz) was injected into the VCSEL and modulated with a 200 ps drop in power.

are exhibited by biological neurons and are desirable for the construction of spiking neural networks can also be demonstrated experimentally in the optical spiking responses of VCSELs, as it will be showcased in upcoming sections of this chapter.

Threshold-and-fire

Thresholding is the observation that the spiking response has a defined boundary between the amplitude of pulses that do trigger spikes v.s. those that do not (for pulses of the same shape and duration), with that boundary known as the threshold. Figure 2.2.5 shows the response of the 1550 nm VCSEL neuron to a sequence of 200 ps long rectangular pulses modulated onto the injection, each with a higher amplitude than the previous. This is used to give a relative measure of the threshold amplitude of one of these pulses: the mean power of the injected light can be adjusted such that several of the higher amplitude pulses trigger spikes, while those below do not. Figure 2.2.5 (a) shows the full traces of the voltage output of the AWG (The MZM was biased at 4.5 V , so the increasing voltage from the AWG causes a drop in injected power), and the photodetector's recording of the VCSELs output, showing the spikes occurring for the highest amplitude pulses. The VCSEL output is wrapped into a temporal map modulo the 2 ns pulse repetition rate in Figure 2.2.5, such that the optical spikes from the VCSEL, marked in bright green, appear at equal times in the map.

For the case shown, a 1550 nm VCSEL was injection locked using $31\text{ }\mu\text{W}$ of 1555.65 nm light (detuned by 2.9 GHz from the subsidiary peak), with a driving current of 4.5 mA . Spikes were then triggered by pulses only higher than 70% of the AWG's voltage range (0.7 V before amplification). Additionally, for the lowest amplitude pulses that could trigger spikes, the time of the spike was delayed compared to those from the strongest pulses. Previously done work by our group [69] has explored more precise measurement and characterising of the threshold for spiking in VCSELs, including for both electrically and optically triggered spikes.

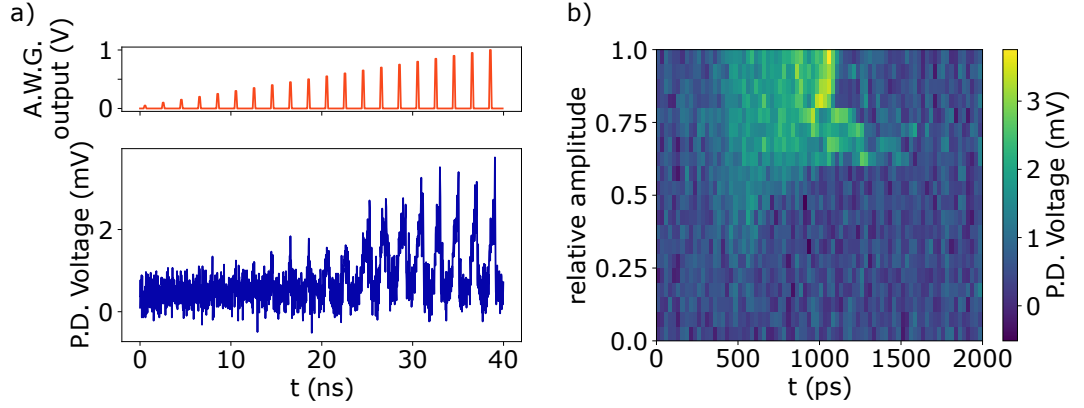


Figure 2.2.5: Threshold-and-fire dynamics of a VCSEL neuron. $31\mu W$ of light was injected at a detuning of $-2.9GHz$ from the subsidiary mode of the $1550nm$ VCSEL. a) The A.W.G. voltage used to modulate the amplitude of the injected light, and the response of the VCSEL. Each input pulse is $200ps$ long. b) Temporal map of the response to a pulse of each amplitude, relative to the range of the A.W.G. (averaged over 10 repetitions). Spikes are only emitted by the VCSEL for pulses higher than 0.7 of the A.W.G. range. Additionally, spikes emitted as a result of pulses that only just cross the threshold (0.7-0.8) are slightly delayed compared to those emitted in response to stronger inputs (0.8).

Integrate-and-Fire

VCSEL neurons also display leaky-integrate-and-fire dynamics: multiple sub-threshold inputs that occur near enough to each other in time may collectively trigger a spike. This behaviour is typical of biological spiking neurons, and is widely used in neuron models in spiking neural networks[99].

This effect was demonstrated using pairs of $200ps$ long rectangular pulses (all of equal amplitude), with the optical injection power and modulation amplitude adjusted such that the VCSEL was injection locked to the externally injected signal, but each single optical input pulse was configured with a sub-threshold amplitude and did not trigger a spike from the VCSEL. The input pulse pairs were then arranged with an increasing time difference between them. Pulse pairs near enough in time together were able to trigger a spike, as shown in Figure 2.2.6, where pairs within $300ps$ of each other were integrated by the VCSEL and triggered a spike, while those that were further apart made only a sub-threshold perturbation.

In Figure 2.2.6, a $1550nm$ VCSEL was in the same state as for the threshold

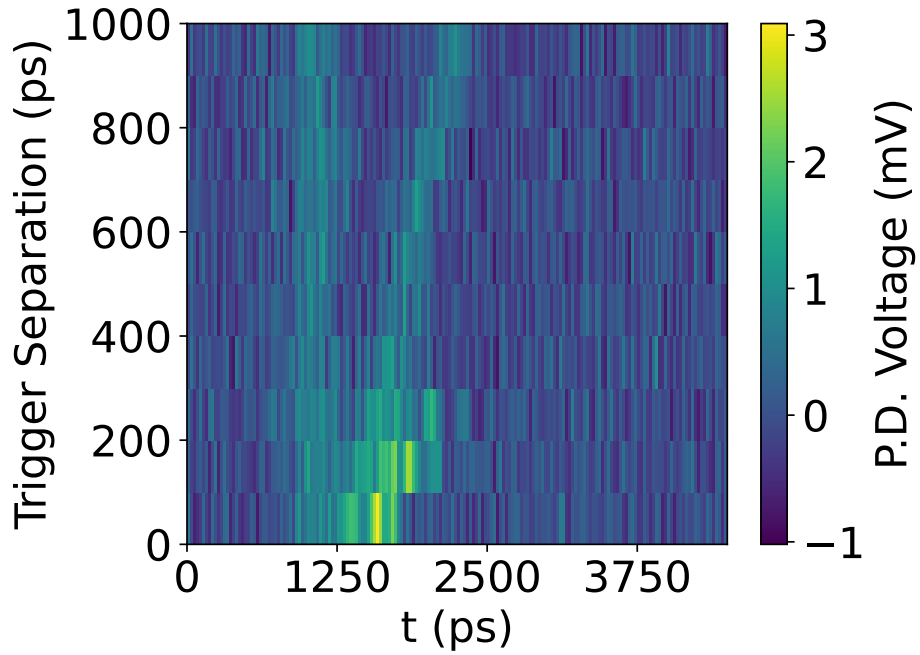


Figure 2.2.6: Integration-and-fire spiking response of the 1550 nm VCSEL used in this work. Two square pulses of 100 ps are modulated on to the injected light, and the amplitude of those pulses is adjusted such that one is insufficient to trigger a spike. When the two pulses follow each other within a sufficiently small window, the VCSEL response to each is 'integrated', and the VCSEL is able to fire a spike. In this case, the VCSEL could spike when the pulses were separated by less than 300 ps , as highlighted by the bright green instances in the map, remaining quiet otherwise.

demonstration, and again injected with $31\ \mu W$ of light, detuned by $-2.9\ GHz$. The modulation used $200\ ps$ rectangular pulses with an amplitude 60% the range of the AWG ($0.6V$ before amplification), these pulses did not trigger spikes, however as shown in Figure 2.2.5, these are near to the amplitude threshold required for spiking.

Further demonstration of the integration effect, involving a wider range of sub-threshold inputs for exploring the precise dynamics of integration can be found in [69].

Refractoriness

Another important dynamical effect characteristic of biological neurons is the refractory period of a spike. When a spike is emitted, a short period of time must elapse before another spike can be emitted, regardless of the input - this time is known as the refractory period of the neuron.

The refractory period of the VCSEL neuron can be found using pairs of optical input pulses at varying temporal separations, however with each input pulse over the amplitude threshold required to trigger a spike from the VCSEL. When the second trigger occurs within the refractory period of the spike from the initial trigger, the second trigger will be unable to trigger a second spike.

Figure 2.2.7 shows the temporal map of the refractory period test of a $1550\ nm$ VCSEL, showing clearly the device's refractory period of approx. $500\ ps$. The pulses were $200\ ps$ drops in optical input power, initially separated by $100\ ps$. The VCSEL was injection locked using $41\ \mu W$ of light, detuned by $-3\ GHz$ from the subsidiary peak. In this state, the maximum amplitude pulses generated by the AWG were far over the threshold required to trigger a spike, and so the first pulse in each pair always caused a spike.

If the second input pulse arrives within the refractory period (e.g. before $500\ ps$), a second spike is not seen, as the VCSEL has not recovered and returned back to the excitable state. When the second pulse is later than $500\ ps$, it is able to trigger a second, well-defined spike, without any influence of the first spike on the shape or time of the second.

The nature of the refractory period of a VCSEL neuron can be controlled by the

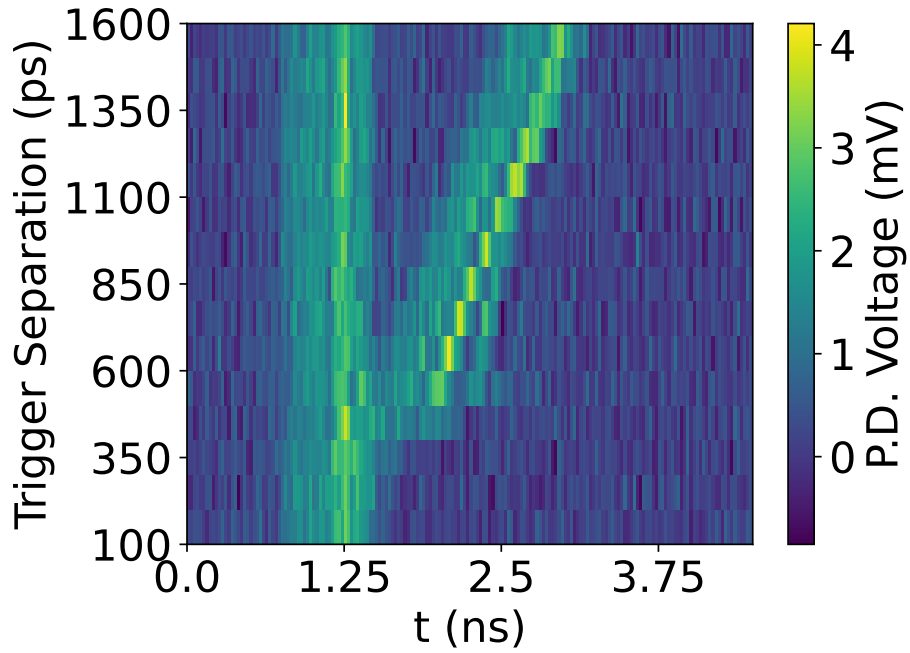


Figure 2.2.7: Demonstration of the refractory period of the 1550 nm VCSEL neuron. As in the case of the integrate-and-fire demonstration, two 100 ps pulses were used to trigger the VCSEL neuron to spike. In this case however, each optical input pulse was strong enough to individually cause an optical spike from the VCSEL. When the two trigger pulses are too close, and the VCSEL has not had time to return to the stable, locked state after the first spike, the second trigger pulse cannot therefore trigger a spike. This shows the VCSELs refractory period of approx. 500 ps for this device.

power of the injected light, with high injection powers leading to a sharply defined cutoff time, while low injection powers result in a 'relative' refractory period, in which the probability of a second trigger causing a spike is reduced, rather than completely eliminated [70, 71].

2.2.3 Spike Modulation

In order to construct larger networks of spiking neurons, it is useful to be able to weight the amplitude of spikes emitted by a VCSEL, or to be able to selectively suppress them altogether. Weighting of optical spikes from VCSELs has been done by our group previously using vertical-cavity semiconductor amplifiers (VCSOAs)[100], and microring resonators[51], and here a Mach-Zehnder modulator at the VCSEL's output is used.

Figure 2.2.8 shows a diagram of the extended experimental setup, including the additional polarisation controller and MZM after the circulator's output, before the final measurement equipment (OSC and OSA). The capability for this setup to both deterministically trigger optical spikes from the VCSEL, and to suppress them at the output is demonstrated in figure 2.2.9: a 1300 nm VCSEL is driven with a bias current of 4 mA , and injected with $107\text{ }\mu\text{W}$ of 1287.86 nm light (detuned by -3 GHz from the subsidiary mode), with a 250 ps drop in the input power being used to trigger an optical spike. At the immediate output of the VCSEL, the light is split by polarisation (as in fig 2.2.4), so that the observable increase in power lasts for long enough to be modulated by the MZM. Trigger pulses were sent to the VCSEL at 2 ns and at 12 ns in fig. 2.2.9, to deterministically trigger two optical spikes.

The voltage modulation sent to the output MZM consists of 1 ns long pulses, such that the MZM allows full transmission during that 1 ns , and minimal otherwise. Two voltage pulses were sent to the output modulator (highlighted in grey in fig 2.2.9), one that was coincident with the optical spike from the VCSEL (at 2 ns in fig. 2.2.9), so that the spike could be seen, and a second was sent at a time when the VCSEL was quiet (at 22 ns in 2.2.9), demonstrating that the final output intensity only increases when the VCSEL has spiked, and not just due to the modulator's transmission increasing.

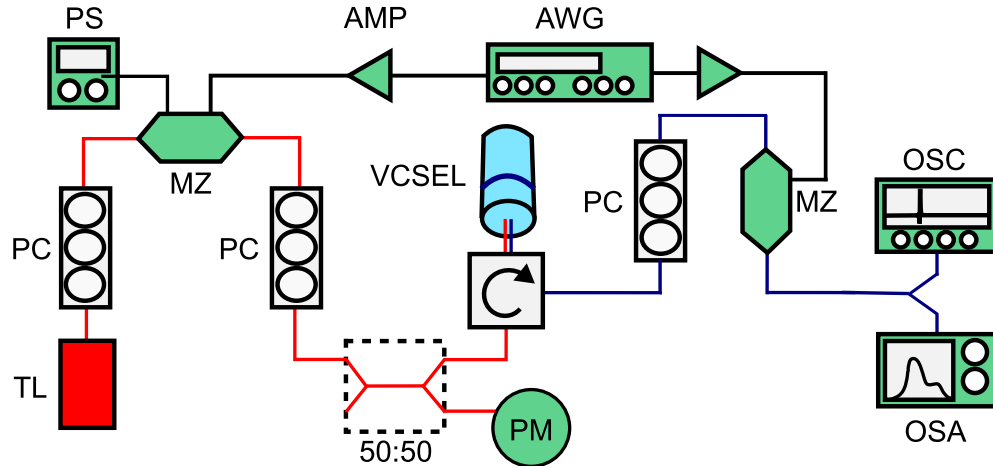


Figure 2.2.8: The experimental setup used to investigate the modulation of optical spikes at the output of a VCSEL neuron. The input optical path follows the same layout, with a tuneable laser providing power that is intensity-modulated using a Mach-Zehnder interferometer and subsequently injected into the VCSEL. In this setup however, the output of the circulator, that is carrying the output of the VCSEL, is fed through a further polarisation controller and modulator. This second modulator is then used to selectively allow or suppress the optical spikes coming from the VCSEL.

2.3 Summary

This section shows the operation of Vertical Cavity Surface Emitting Lasers, and their implementation as photonic spiking neurons.

The nonlinear dynamics exhibited by VCSELs under optical injection were investigated experimentally, showing that near the boundary of injection locking, carefully controlled perturbations can drive the laser into an excitable regime in which it produces ultrafast optical spikes. These spike responses arise from polarization switching dynamics and exhibit neuron-like properties including thresholding, temporal integration, and refractoriness.

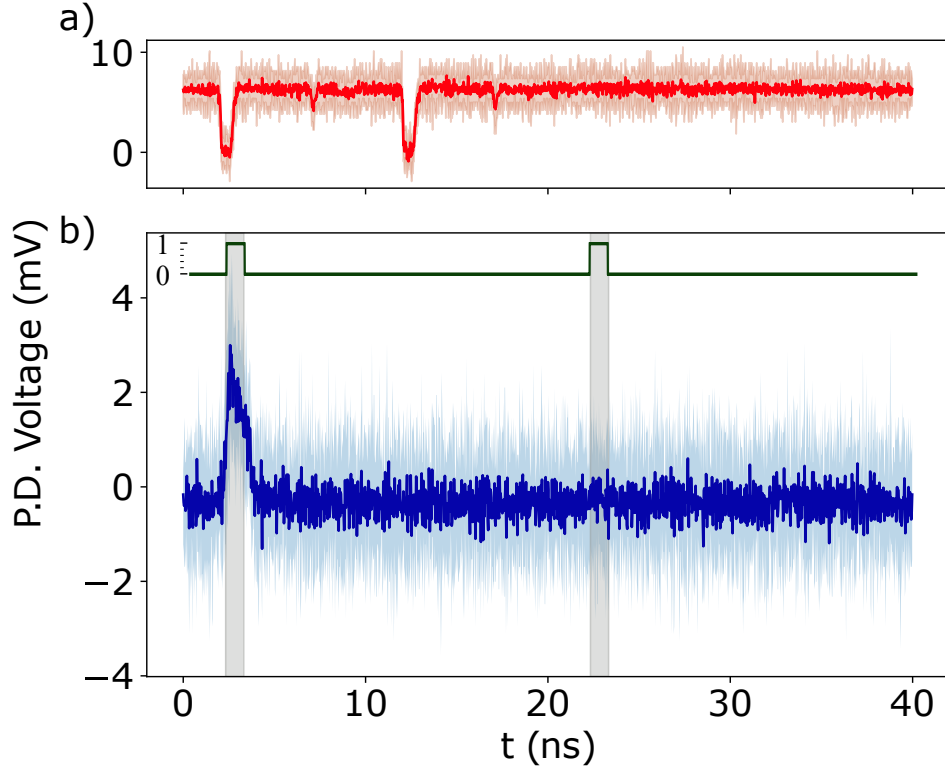


Figure 2.2.9: Combining modulation at the input and output of a VCSEL. A first modulator is used to create a spike-triggering drop in the injection power, while a second modulator is used at the VCSEL's output to selectively allow or block the spike. For this case, the VCSEL output is polarisation-split, so that the longer increase-only component is used, to allow for jitter in the timing of the output modulation relative to the input (and subsequent spike). a) The amplitude of light injected into the VCSEL. Drop triggers occur at 2 ns and 12 ns , which each trigger a spike from the VCSEL. b) The output of the VCSEL after polarisation splitting and modulation. The output modulator (modulation plotted by the green line) is controlled to allow maximum transmission for 1 ns at time instants 2 ns and 22 ns (grey shaded regions), and is set to minimum transmission at all other times. The result is that the spike that occurs at 2 ns is visible, while the other is suppressed by the modulator. The output modulator on its own (22 ns) also does not cause a notable increase in the recorded power.

Chapter 3

Photonic Neural Networks Based on VCSELs

Building on the foundation of VCSEL spiking neurons, this chapter presents approaches for constructing neural network architectures using VCSELs, with an emphasis on efficient and scalable implementations. Different strategies for encoding input data, generating dynamic responses, and extracting computational features are outlined, highlighting how temporal dynamics can be used as a resource for computation. The chapter also considers practical aspects of system design, including methods for increasing effective network size and reducing hardware complexity.

The capabilities of VCSEL-based neural networks are demonstrated through a series of representative computational tasks, illustrating their potential for fast and efficient processing. Through these examples, the chapter provides an overview of how VCSELs can be used to realise photonic neural networks, and positions them as a promising approach within the broader field of neuromorphic photonics.

Section 3.1 introduces the architecture and experimental setup of the VCSEL neural network, as will be used throughout the chapter, and demonstrates its function on a standard machine learning task. The following section describes a novel algorithm for training the output weights of this neural network, leveraging the discrete nature of the spiking response, and section 3.3 extends the use of the network to time-series

prediction, a task involving continuous values - unlike the previous tasks.

Relevant Publications

- D. Owen-Newns, J. Robertson, M. Hejda, and A. Hurtado, “GHz Rate Neuromorphic Photonic Spiking Neural Network With a Single Vertical-Cavity Surface-Emitting Laser (VCSEL),” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 29, no. 2: Optical Computing, pp. 1–10, Mar. 2023, issn: 1077-260X. doi: 10.1109/JSTQE.2022.3205716.
- D. Owen-Newns, J. Robertson, M. Hejda, and A. Hurtado, “Photonic spiking neural networks with highly efficient training protocols for ultrafast neuromorphic computing systems,” *Intelligent Computing*, vol. 2, p. 0031, 2023. doi: 10.34133/icomputing.0031.
- D. Owen-Newns, L. Jaurigue, J. Robertson, A. Adair, Andrew, J. Jaurigue, K. Lüdge, and A. Hurtado “Photonic spiking neural network built with a single vcsel for high-speed time-series prediction,” *Communications physics*, vol. 8, no. 1, pp. 110–9, 2025, issn: 2399-3650.

3.1 VCSEL-based Photonic Spiking Neural Network

In this section, we introduce the design and experimental realisation of the VCSEL-based photonic Spiking Neural Network (p-SNN). The network makes use of time-division multiplexing (TDM), in which the nodes of the network are represented by a measured quantity (e.g. the intensity output of a laser) at chosen intervals of time (as opposed to spatial multiplexing, where the nodes are spatially distributed but simultaneous in time).

Previous work with VCSELs has been based on analog (not spiking) TDM reservoir computers (RC)[44, 70, 101], in which a laser is connected to itself via an optical feedback loop, creating a recurrent neural network structure. The virtual (time-multiplexed) nodes of these networks are linked by the highly nonlinear dynamics of the VCSEL under delayed optical feedback. Training a reservoir computer is much simpler

than training a typical deep neural network (particularly recurrent neural networks) as the only trained parameters are the weights of the final 'output' layer of the network - all other 'weights' are random and implemented physically by the dynamics of the laser-feedback system.

This section introduces a VCSEL-based p-SNN that operates similarly to a VCSEL reservoir computer - using time-multiplexed virtual nodes, and the nonlinear dynamics of the laser, however without the optical feedback line, running purely as a feedforward system. As shown previously, VCSELs can operate as artificial spiking neurons as they exhibit nonlinear neural-like behaviours (e.g. leaky-integrate and fire responses and refractoriness), and these excitable spiking dynamics are utilised to form the network architecture.

In VCSEL-based spiking TDM systems [46], the neuron-like responses in spiking VCSELs couple multiple time-multiplexed virtual nodes for the creation of complex, interconnected p-SNN structures. The coupling of virtual nodes in this way is possible if the temporal duration of the virtual nodes (θ) is set shorter than the time-scale of the neuronal spiking non-linearity. Importantly, the coupled time-multiplexed nodes now deliver spike firing responses instead of a continuous output, using neural-like signals for operation. Using this technique (depicted graphically in Fig. 3.1.1 (b)), a single VCSEL can deliver a system of optically coupled spiking virtual nodes (with sub-nanosecond temporal operation) yielding a novel photonic SNN architecture. Unlike previous VCSEL-based RC reports [46], this system (shown in Fig. 3.1.1 (b)) has no recurrent/feedback connections, thus only requires the hardware implementation of a single spiking VCSEL. In this approach, the input signals entering the VCSEL-based p-SNN are chosen to be analogous to those used previously in traditional non-spiking laser-based photonic RC systems (i.e. continuously modulated light input) [46].

Differing from photonic RCs, the signals used for computation (at the output of the VCSEL p-SNN) are event-based (sub-nanosecond long optical spikes) and binary (as seen in Fig. 3.1.1 (b)): the value of the system's output nodes are directly determined by the firing (binary '1') or non-firing (binary '0') of a short 150 ps optical spike within the timeframe θ of the virtual node. This removes the down-sampling or smoothing

operations often performed during continuous value readouts in other non-spiking RC systems. Further, the lack of feedback grants additional stability to the p-SNN system, as well as performance flexibility, with the number of virtual nodes selected without consideration of hardware manipulation.

The experimental setup used to build the VCSEL-based p-SNN is shown in Fig. 3.1.1 (a). This all-optical fibre setup consists of a 1300 nm tuneable laser (TL) whose output is externally modulated by a Mach-Zehnder (MZ) intensity modulator. The modulated TL light's is subsequently optically injected into a 1300 nm VCSEL operating as the spiking photonic neuron. The optical injection power is measured using a power meter (PM) and the corresponding output of the VCSEL-based p-SNN and spiking RC system is measured using an optical spectrum analyser (OSA), a fast ($> 9\text{ GHz}$) amplified photodetector and a high-speed (16 GHz) real-time oscilloscope (OSC). An arbitrary waveform generator (AWG) operating at 12 GSa/s generated the input waveforms to the spiking RC system. The output signals of the AWG are amplified with a 14 GHz electrical amplifier (AMP) before their injection into the RF input of the MZ modulator, which encodes the input waveforms into the TL's light. The DC bias input of the MZM is set using a power supply (PS). The 1300 nm VCSEL used in the experiments had a measured lasing threshold current of 1.4 mA at room temperature (293 K). The VCSEL's spectrum showed two peaks, corresponding to the two orthogonal polarisations of the fundamental transverse mode of the device (see [101] for exemplar device spectral characteristics). The polarisation of the TL's injected light is matched to that of the subsidiary attenuated mode of the VCSEL (using Polarisation Controllers, PCs) and injected such that the VCSEL injection-locks to the TL's light emission.

3.1.1 Nonlinear Classification Task Operation

The task selected to demonstrate the performance of the new VCSEL-based photonic spiking system is Fisher's 1936 iris flower dataset, an established benchmark Machine Learning classification task [102]. Each data point in this dataset (plotted in Fig. 3.1.2) has 4 variables, representing the dimensions of particular features (the sepal and petal

lengths and widths) of multiple Iris flower specimens, and includes a total of three iris species/classes (Setosa, Virginica & Versicolor). The dataset is made up of 150 different flower specimens (50 of each species), and the task is to identify which species a flower is given its set of measurements. Crucially, only one of the species (Iris Setosa) is separable from the others in this feature space by a plane (a linear discriminant), while the other two (Iris Virginica and Versicolor) are separated from each other by a nonlinear, curved boundary. Using the 4 variables, namely Sepal Length (SL), Sepal Width (SW), Petal Length (PL) and Petal Width (PW), as inputs for the spiking RC system, we can effectively create a p-SNN architecture (depicted in Fig. 3.1.3) capable of classifying the 150 Iris flowers into their respective 3 species.

Here, we make use of time division multiplexing to create the virtual spiking nodes in the VCSEL-based p-SNN. Sampling the temporal spiking output of the VCSEL at time intervals of θ , provides the spiking or non-spiking output state of each node in the p-SNN. These nodes are interconnected by the system's neural-like dynamics (e.g. leaky integrate-and-fire behaviour, refractory period), which occur on timescales around 1 ns. Hence, using a spiking node duration θ lower than 1 ns allows neighbouring nodes to influence one another, creating fixed network connections and therefore forming an SNN architecture with a single VCSEL system. This also allows the number of spiking nodes (N_v) to be modified at will, by simply setting the temporal length and characteristics of the input data waveforms injected into the VCSEL-based photonic spiking neural network.

To prepare the data for injection into the p-SNN, first a random input weight matrix is created with 4 columns and as many rows as spiking nodes (N_v) in the p-SNN. This matrix multiplies a column vector whose components are the 4 Iris flower data point variables (SL, SW, PL, PW), giving a final vector (with N_v components). The components of this vector gives the inputs that are generated by the AWG, encoded in the TL's light output by the MZ modulator, and optically-injected into the photonic VCSEL-based spiking RC system. The total input waveform is composed of all 150 flower data points, concatenated in series (as shown in Fig. 3.1.4(a)), with all Setosa Iris flower species directly followed by the Versicolor and Virginica species. Figure

3.1.4(b) shows an expanded section of the input waveform in Fig. 3.1.4(a), which represents a single input data point (flower number 1) for the photonic spiking system.

Periods of no input modulation (lasting 2 ns , longer than the integration or refractory times of the VCSEL) are added between each data point to allow the VCSEL to reach quiet equilibrium between consecutive data inputs. This ensures that the response of the VCSEL is only due to the input generated by the relevant data point, and is not influenced by previous data points, removing any potential bias that may otherwise be introduced by sequential data points. These can be seen as the minima values in the input waveform of Fig. 3.1.4(a), and more clearly after the input data for flower 1 (at approx. p-SNN Input $y = 2.85 \times 10^3$, time = 550 ns).

The optical output of the VCSEL, representing the output states of all the spiking nodes in the photonic SNN, is recorded with a real-time oscilloscope (OSC). Figure 3.1.4(c) shows the measured optical output from the VCSEL for the same time segment of input data (flower number 1) included in Fig. 3.1.4(b). The output from the VCSEL takes the form of a fast optical spiking pattern (see Fig. 3.1.4(c)), showing multiple excitable neuron-like spikes occurring at the output of the different virtual nodes of the p-SNN architecture. To detect which nodes have fired spiking responses the recorded output is cut into bins of θ (such that there are N_v bins). If the peak voltage in a bin exceeds the chosen spike threshold, that bin is taken to contain a spike. This process creates a vector of binary (spike & no-spike) values that are used to produce, using linear-least-squares fitting, the weight matrix of the output layer of the photonic SNN.

3.1.2 Experimental Performance of the Photonic Spiking Neural Network

The VCSEL-based p-SNN was configured to run with a spiking virtual node duration of $\theta = 250\text{ ps}$, and two different configurations, with $N_v = 512$ and $N_v = 1024$, were tested.

When running the system with 512 spiking nodes, the VCSEL was driven with a bias current of 4.02 mA (approx. 2.9 times the device's lasing threshold current) and at a constant temperature of 293 K . The optical injection from the TL was made with

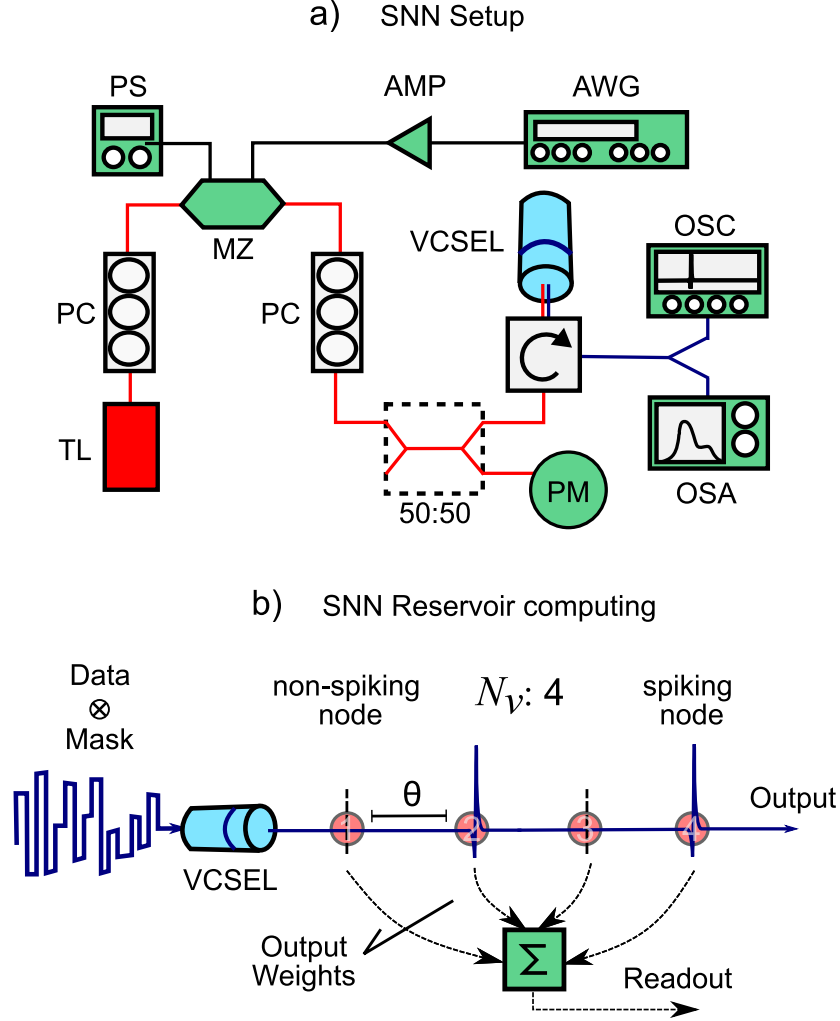


Figure 3.1.1: Experimental setup and p-SNN operation. a) The experimental setup for the VCSEL-based spiking reservoir computer/SNN. An external Tuneable Laser (TL) injects intensity-modulated light into the VCSEL via an optical coupler and circulator. Injection light polarisation is controlled by polarisation controllers (PCs). The masked input signals are generated by an arbitrary waveform generator (AWG) and amplified electrically (AMP) before entering the DC power supply (PS) -controlled Mach-Zehnder intensity modulator (MZ). The optical spiking signals produced by the VCSEL are analysed using a 16 GHz oscilloscope (OSC) and an optical spectrum analyser (OSA). b) The operation of the spiking photonic RC system. Input data is multiplied by a random mask before injection into the VCSEL. The VCSEL is operated near the injection locking boundary where inputs can produce non-linear spiking responses. Inputs occurring within the refractory period/non-linear timescale of the spiking dynamics ($\sim 1\text{ ns}$) are internally connected, hence producing unique spiking patterns at the p-SNN output. The virtual nodes are measured from the optical output of the VCSEL at intervals of θ ($\theta = 250\text{ ps}$), internally connecting 4 sequential virtual nodes. The output weights are applied to each node and linearly combined to provide the readout of the reservoir.

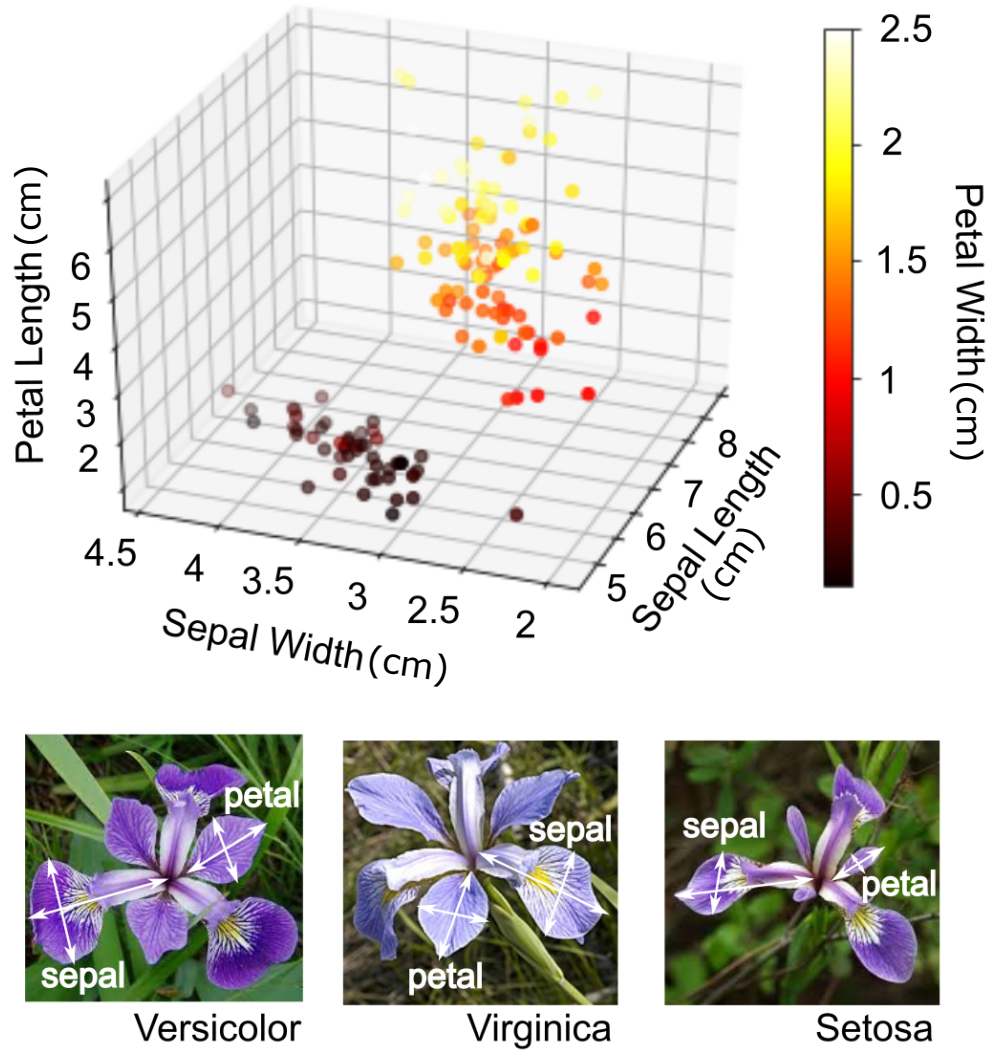


Figure 3.1.2: Scatter plot of Fisher's Iris data used in the classification task [103]. One class (Setosa) is separable from the other two (Versicolor & Virginica) by a plane, and as such is separable by linear transform. The Versicolor & Virginica classes however cannot, and the separating surface is more complex, but can be learned by a neural network.

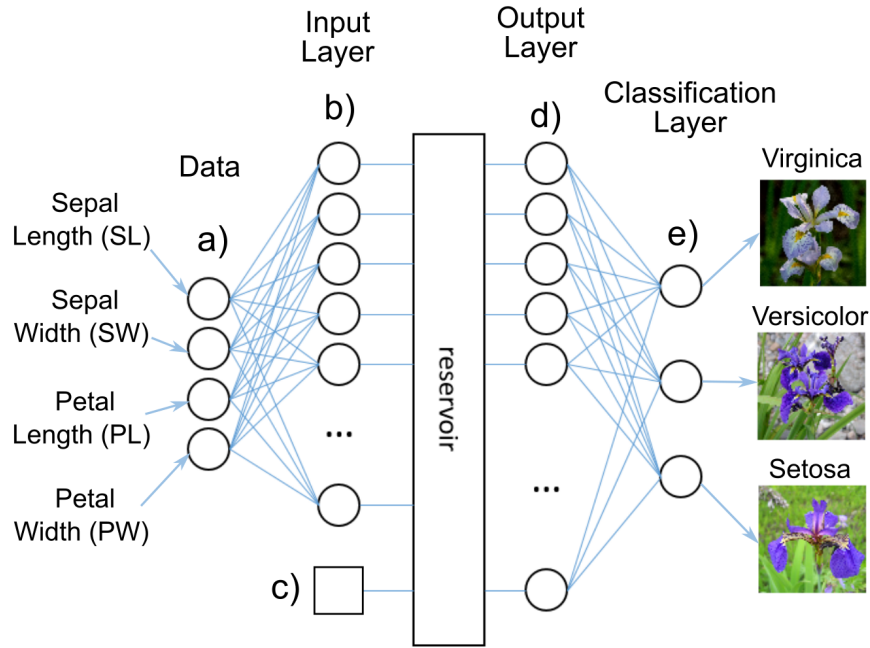


Figure 3.1.3: The logical architecture of virtual nodes in the p-SNN. Each node is distributed in time, separated by θ . The reservoir represents the nonlinear coupling between each of the input nodes. a) The 4 data variables (SL, SW, PL, PW). b) The input layer where values are calculated by applying a random mask to the data. c) A disconnected input representing the time-spacing introduced between data points. d) The output layer where binary values represent whether a spike has or has not occurred. e) The classification layer that provides the probability a data point belongs in each class. The total number of nodes in the both the input and output layer of the p-SNN is equal to N_v . The only weights that require adjusting (training) are the connections between d) and e).

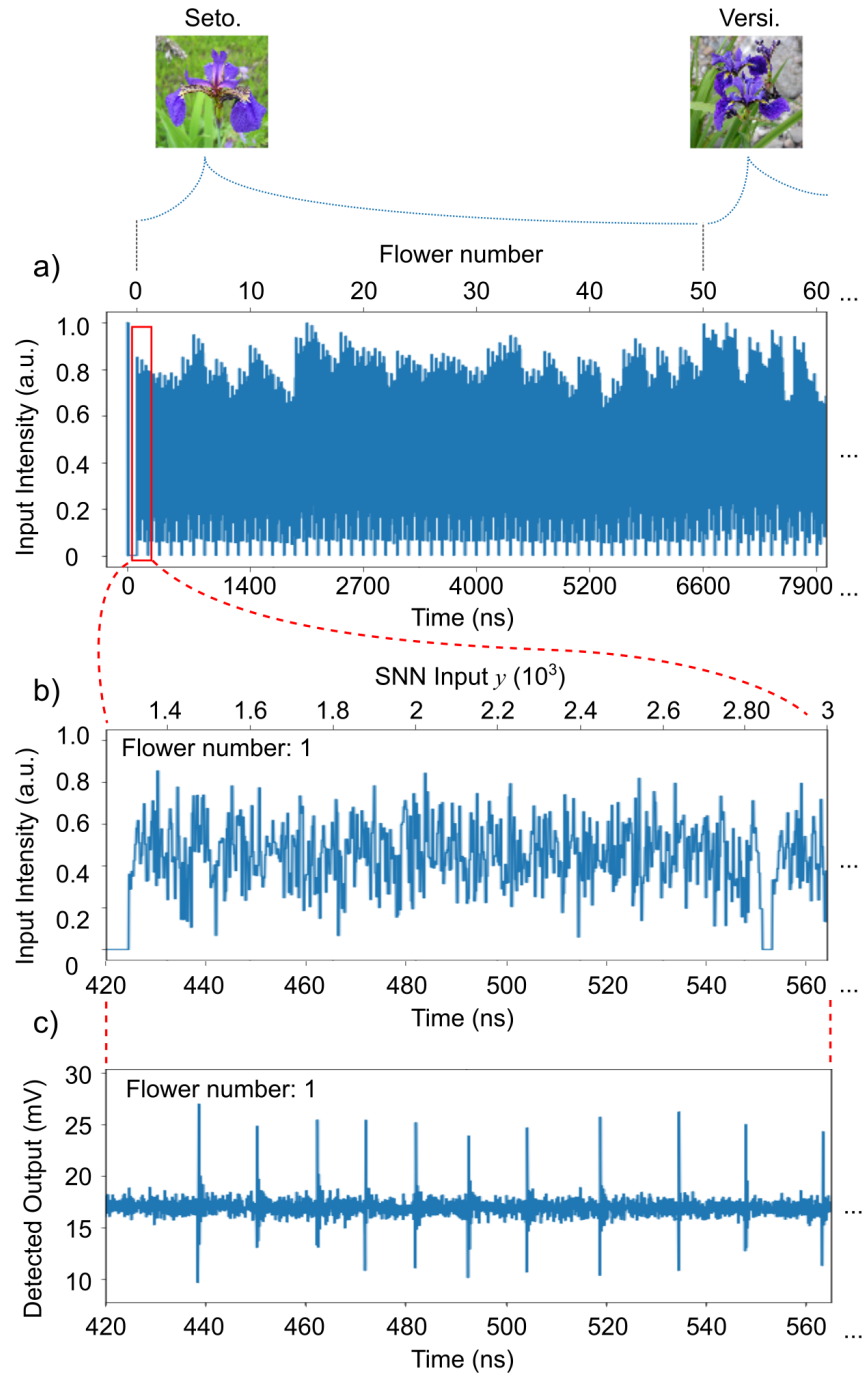


Figure 3.1.4: (a) Section of time-series containing the prepared p-SNN input of all 150 masked data points. The input data points are concatenated by species and separated by 2 ns . Input shown is for 512 virtual nodes. b) Highlighted red section containing the p-SNN input of flower number 1. p-SNN input has been resampled for AWG generation (12 GSa/s - 3 samples per θ). c) The recorded output time-series of the p-SNN, showcasing the spiking pattern obtained for flower number 1.

181 μW average optical power and an initial frequency detuning of $-1.1 GHz$ between the externally-injected optical signal and the resonance of the VCSEL’s subsidiary polarisation mode. For the system test with 1024 spiking nodes, the driving bias current applied to the VCSEL was equal now to 3.92 mA (approx. 2.8 times the lasing threshold current), and the optical injection was performed with a 159.4 μW average optical input power at a frequency detuning of $-3 GHz$ between the resonant frequency of the TL’s light and the VCSEL’s subsidiary attenuated mode. Near-identical spiking regimes were obtained using each set of experimental conditions, allowing for consistent photonic SNN operation on different experimental runs. The stability of the system against changes in environment (e.g. ambient temperature) and system parameters was determined by the stability of the spiking regime, which is described in previous works[101].

Figure 3.1.5 plots the experimentally measured temporal maps of all the optical spike patterns produced at the output of the VCSEL-based p-SNN in response to all 150 data points in the Fisher’s Iris flower dataset. Temporal maps are plotted for both tests of 512 (Fig. 3.1.5 (a)) and 1024 (Fig. 3.1.5 (b)) spiking nodes. In the temporal maps of Fig. 3.1.5, the nodes that displayed (did not display) a spike firing event are plotted in green (black). The temporal maps show that for both cases investigated each data point (each input flower specimen) produces a characteristic response with distinct spiking patterns shared across flowers of the same species (class). The results in Fig. 8 show that specific spiking patterns arise in the system’s response to each flower species, with selected nodes firing and others remaining silent consistently.

By interpreting the output state of each spiking (virtual) node in the photonic SNN, the weights of the output layer were trained using a linear-least-squares fit. The data used for training was arranged into two matrices: a first matrix in which the rows were the spiking patterns produced by each data point, and a second matrix with three columns to label the data (i.e. a value of ‘1’ in the first, second or third column denotes the data point as classes ‘1’, ‘2’, and ‘3’, respectively). The output layer weight matrix is found by linear-least squares fitting the spike vectors S to the labels L . In S the virtual nodes with and without elicited optical spikes are represented by binary ‘1’ and

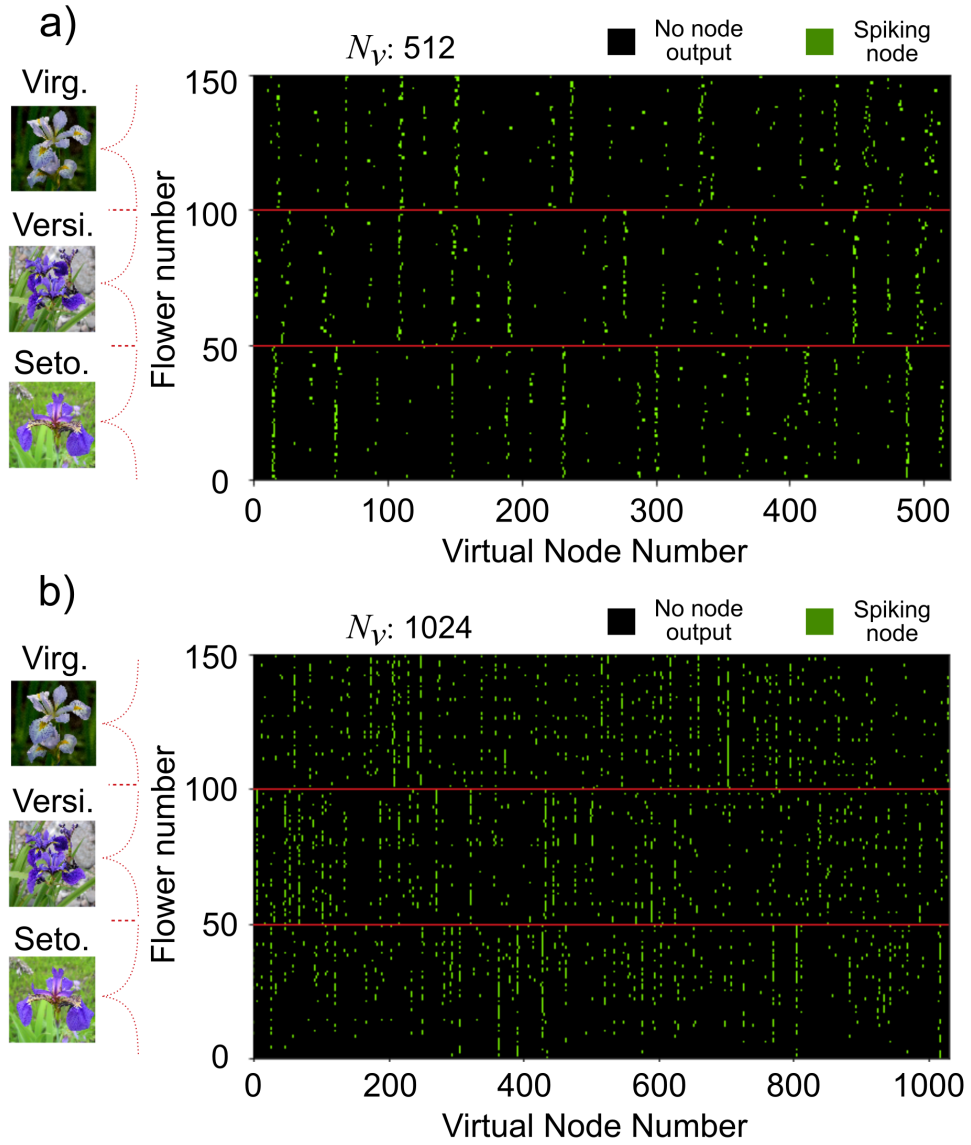


Figure 3.1.5: Spike patterns produced by the p-SNN for all 150 flower data points (50 of each species). The system was tested using (a) 512 and (b) 1024 virtual nodes, with each input flower producing a different spiking pattern. Green represents the presence of a spike within the virtual node.

‘0’ values, respectively. The weight matrix can be calculated directly from the product of the Moore-Penrose inverse of S and L , giving a matrix with N_v rows and 3 columns ($W = S^+ * L$). Subsequently, by multiplying the output weight matrix W by the spiking pattern produced by the p-SNN in response to a new input data point, we produce a triplet containing the estimate of the likelihood that the data point belongs to each of the three Iris flower classes. The class with the highest likelihood is then taken as the predicted class of that data point.

Training was done using an equal random subset of data points from each of the three data classes, with the remaining points used to test the accuracy of the VCSEL p-SNN. Figure 3.1.6 provides confusion matrices obtained from the VCSEL p-SNN (with 512 or 1024 spiking nodes) when performing the Iris flower classification task. The confusion matrices in Fig. 3.1.6 were obtained for a training set size of 30 (10 of each flower class) and a test set size of 120 (40 of each flower). Fig. 3.1.6(a) shows that the p-SNN correctly classifies almost all of the flower data points yielding an average classification accuracy (correct classifications divided by the number of test points) of 0.917. Fig. 3.1.6(b) shows that once the number of nodes in the p-SNN is increased to 1024, the p-SNN correctly classifies all 120 tested data points, reaching full accuracy.

Further, we investigate the performance of the VCSEL-based p-SNN when different training set sizes are applied to both configurations of 512 and 1024 virtual spiking nodes. The curves plotted in Fig. 3.1.7 provide the system’s classification error obtained across 10 different experimental runs for both cases of 512 and 1024 spiking nodes. The calculated average error in each of the two cases is given by the dark blue solid line (see Fig. 3.1.7). For both cases, Fig. 3.1.7 shows that as expected, at first for very small training dataset sizes (up to 3 points) the system’s error is very high. However, remarkably, the system’s performance improves drastically in both cases for very small training set sizes (from values as low as >5 training points), reaching classification error minima at around training set sizes of 10 (with larger training sets, the test set is diminished and so the precision in the error decreases). In both cases classification accuracies remaining consistently over 90% and 97% when the VCSEL p-SNN is configured with 512 or 1024 spiking nodes (node numbers outside this range are not

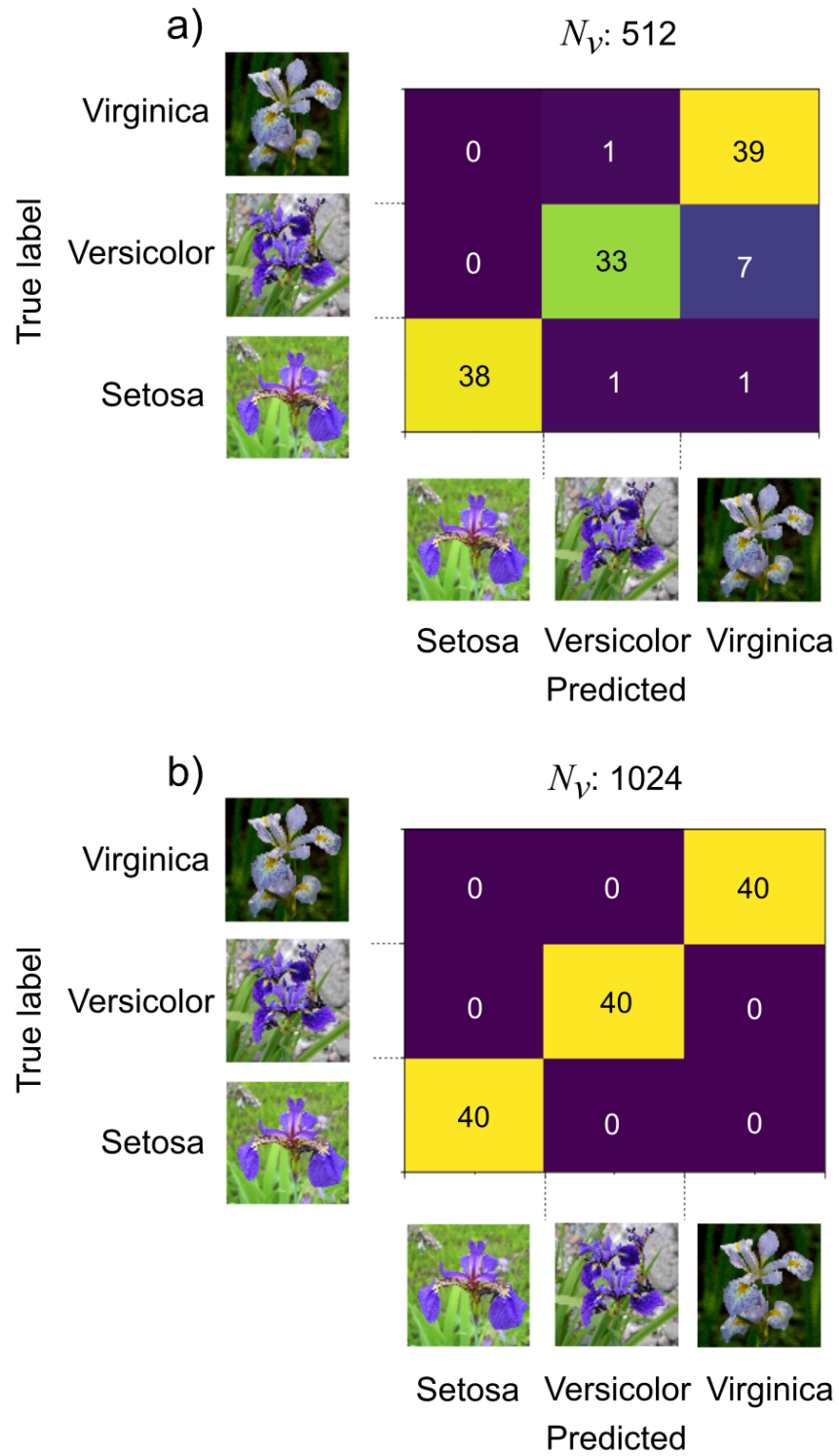


Figure 3.1.6: Confusion matrix for the Iris flower classification task using a) 512, and b) 1024 spiking (virtual) nodes in the VCSEL-based p-SNN implementing the photonic spiking RC system. A training set of 10 data points was used for each class.

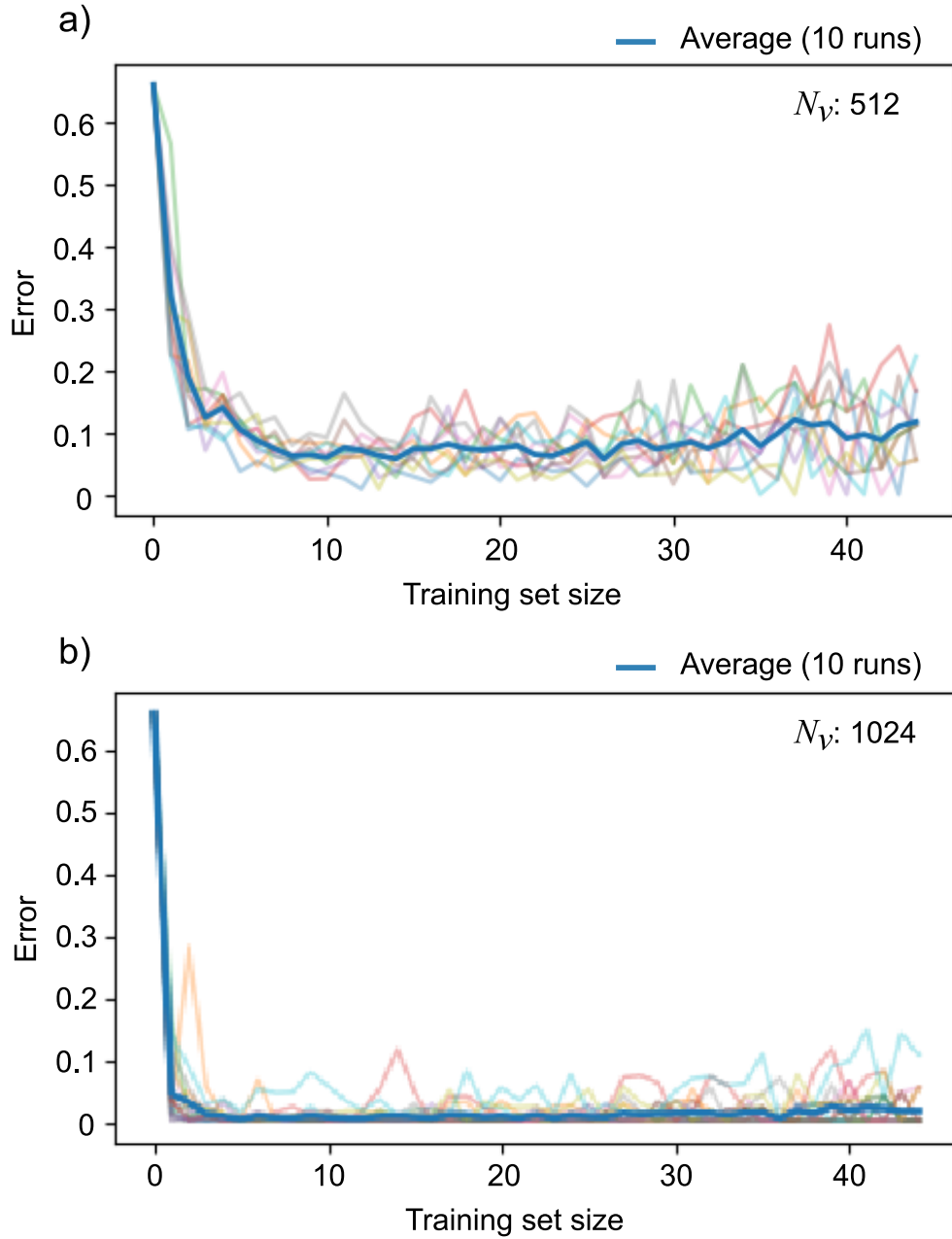


Figure 3.1.7: The classification error is plotted for various training set sizes. The error for runs of both (a) 512 and (b) 1024 spiking (virtual) nodes are shown. Faint lines represent several experimental runs of the VCSEL-based photonic SNN, the dark blue solid line represents the average error. This error is calculated as the number of points incorrectly classified out of the total number of points tested.

particularly useful for the specific task: reducing the nodes number below 512, results in the attainment of lower accuracy levels, whilst increasing the node count above 1024 does not improve the very high accuracy already achieved).

3.2 A Novel Training Method for the Spiking VCSEL p-SNN

This section further demonstrates the capability of the VCSEL p-SNN by applying it to a more complex classification task, and introduces a novel training method tailored specifically to use the discrete spiking response of the VCSEL.

The task used in this second study is an artificial dataset named 'MADELON,' which was created to test feature selection methods as part of a NeurIPS 2003 challenge [104]. The dataset comprises 32 clusters of datapoints placed on the vertices of a 5-dimensional hypercube. Two classes of data are created by randomly labeling each datapoint with the value -1 or +1. Each data point had five informative features (one for each dimension) that could be used to identify a data class. Using these five informative features, 15 linear combinations were created, increasing the feature set size to 20. Finally, 20 features were randomly scattered among 480 non-informative (no predictive power) features, creating a total of 500 features per data point. The task is highly complex, given that a classification must be made using a feature set in which no feature by itself is informative.

The MADELON dataset contains 2000 data points intended for training. However, to demonstrate the VCSEL-based photonic SNN, 300 data points (150 for each class) were used. Limiting the number of input points reduced the length of the time-multiplexed output generated during the experimental runs, thereby allowing us to test different network architectures with diverse node numbers while remaining within the memory of our experimental equipment. As the MADELON task is of much higher complexity than the previous iris flower classification task, a higher number of virtual nodes (N_v) were selected, namely 2048 and 4096, to test the classification performance of the photonic SNN.

The p-SNN architecture used is identical to the architecture used for the flower task: Each of the 500 features was fed into the N_v input layer nodes via a random mask. The input data were masked by multiplying a column vector of all 500 features by a $500 \times N_v$ random matrix, resulting in a vector of N_v components. Experimentally, randomly masked data were generated by the AWG and modulated by the MZM for optical injection of the VCSEL. The output layer of N_v nodes is then provided by the segmentation of the VCSEL spiking output time-series at θ intervals. Again, a node separation (θ) of 250 ps was used for the time-multiplexed virtual nodes, creating multiple interacting nodes within the 1 ns timescale of nonlinearity in the VCSEL. The final classification layer of the p-SNN then provides a -1 or 1 label to the data following the offline training and weighting of the output nodes.

3.2.1 Experimental Procedure

First, the MADELON task was executed in a photonic VCSEL-based p-SNN using a network node count of $N_v = 2048$. This resulted in a total processing time of 512 ns per datapoint in the photonic SNN. The VCSEL was driven with a current of 3.5 mA (~ 2.5 times the lasing threshold current) at a temperature of 293 K and the injected light had a mean power of 142.6 μW and was injected into the VCSEL with an initial frequency detuning of -5.4 GHz with respect to the resonant wavelength of the VCSEL's subsidiary polarisation mode peak.

In the second experimental run, the number of virtual (spiking) nodes in the network was increased to $N_v = 4096$ (1024 ns of processing time per datapoint). In the second run, the VCSEL was biased with a current of 3.45 mA at a set temperature of 293 K. The injected light signal, encoded with the input data, had a mean optical power of 161 μW , and was injected with an initial frequency detuning of -4.6 GHz (with respect to the resonance of the VCSEL's subsidiary polarisation mode).

Figure 3.2.1 presents the experimentally measured time-series depicting the optical input (Fig. 3.2.1 (a-b)) and output signals (Fig. 3.2.1 (c)) of the photonic SNN. The example shown in Fig. 3.2.1 is for a network architecture with 2,048 nodes. Time multiplexing was used to inject each masked datapoint into all (virtual) nodes of the photonic

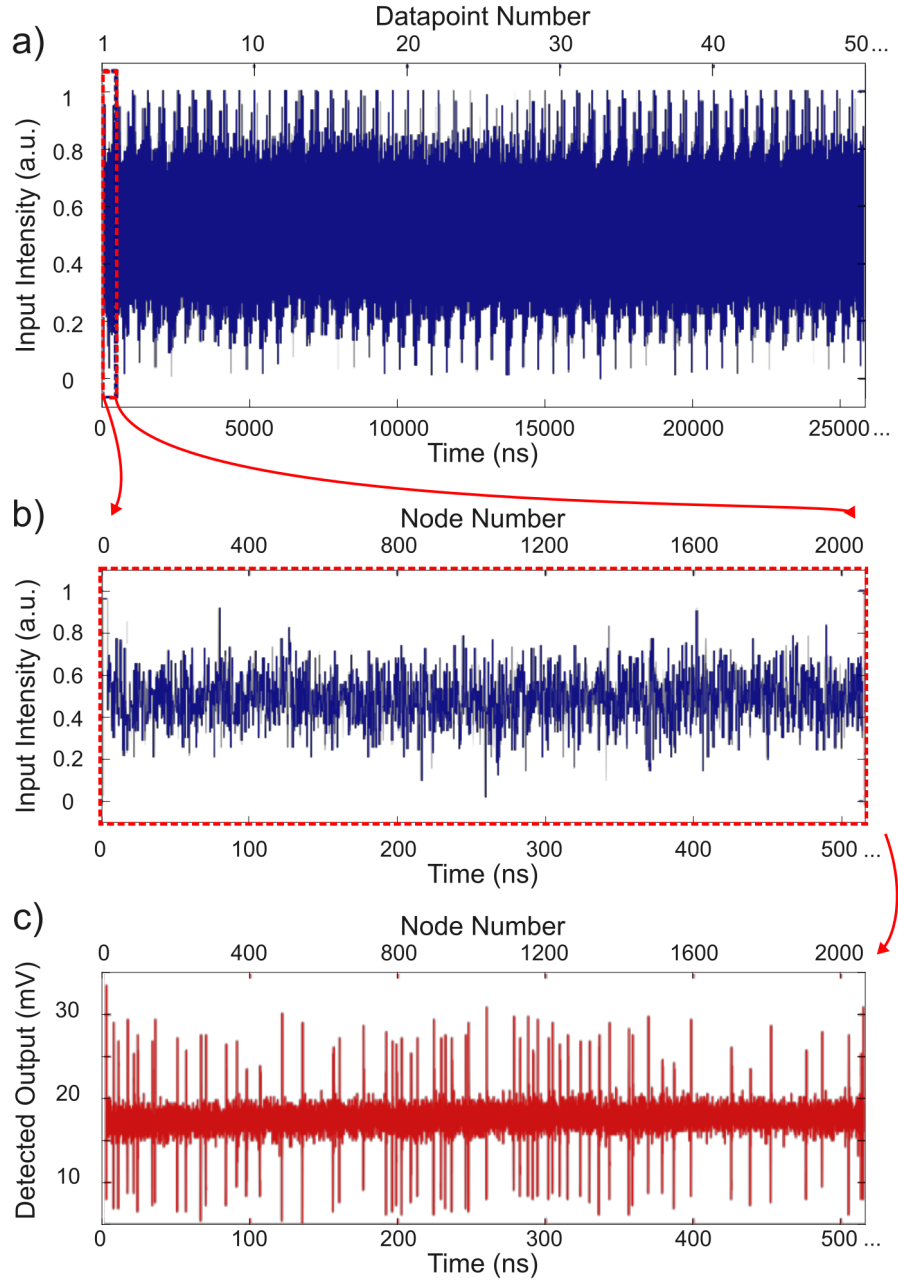


Figure 3.2.1: time-series showing the measured input (blue) and output (red) of the photonic SNN for a network architecture of 2,048 virtual nodes. a) A section of injection waveform showing the time-multiplexed intensity variations of 50 consecutive masked datapoints of the MADELON dataset. b) A 512 ns segment of the input time-series corresponding to the first MADELON datapoint. Each of the 2,048 virtual nodes has a duration of 250 ps, creating the 512 ns input time-series. c) The corresponding measurement of the p-SNN output when injected with the first MADELON datapoint.

SNN (at a rate of 250 ps per virtual node). This created a time-varying signal for each datapoint, which was applied as the intensity modulation of the light injected into the p-SNN. During the experiment, all 300 masked datapoints (time-varying signals) were injected into the p-SNN sequentially, creating a large optical input time-series. Figure 3.2.1 (a) depicts a part of the large optical input time-series, showing 50 sequentially encoded masked datapoints from the MADELON dataset. Figure 3.2.1 (b) shows the first 512 ns of the optical input time-series in Fig. 3.2.1 (a), which corresponds to the input signal encoding the first MADELON datapoint. The optical spiking time-series measured at the output of the p-SNN in response to the first injected MADELON dataset point (input signal in Fig. 3b) is shown in Fig. 3.2.1 (c).

The output of the p-SNN was collected using the same procedure introduced for the Iris classification task: the photodetector voltage reading for each datapoint was cut into the virtual nodes, and each virtual node was assigned a value of 1 or 0, representing the presence (or lack thereof) of a spike. Figure 3.2.2 shows a 2D temporal map that collects in one plot the spiking output from all virtual nodes of the photonic SNN for all 300 consecutively injected data points in the MADELON task. In Fig. 3.2.2 a yellow (blue) dot indicates a spiking (non-spiking) response from the VCSEL-based p-SNN, and the resulting 1 (0) binary output of the virtual node in vector S . The output vector S was plotted horizontally for each input datapoint, revealing the trends in node activation across inputs and classes. The results for p-SNNs configured with 2,048 and 4,096 virtual nodes are plotted in Figures 3.2.2 (a) and (b), respectively. The red line through the center of the map separates the results for each distinct class (-1 & +1) of the input datapoint. Figures 3.2.2 (a) and (b) show that each input datapoint of the same class elicits a similar spike train at the output of the p-SNN. These spike trains were sufficiently distinct from those obtained from other classes of datapoints to allow for a classification operation.

For the p-SNN with a 2,048-node architecture (Fig. 3.2.2 (a)), a total of 298 datapoints (149 per class) were used in the experimental run. This is because the first datapoint in class -1 was affected by an experimental artifact and was removed from the analysis. Therefore, one point from class $+1$ was removed to keep the size of each

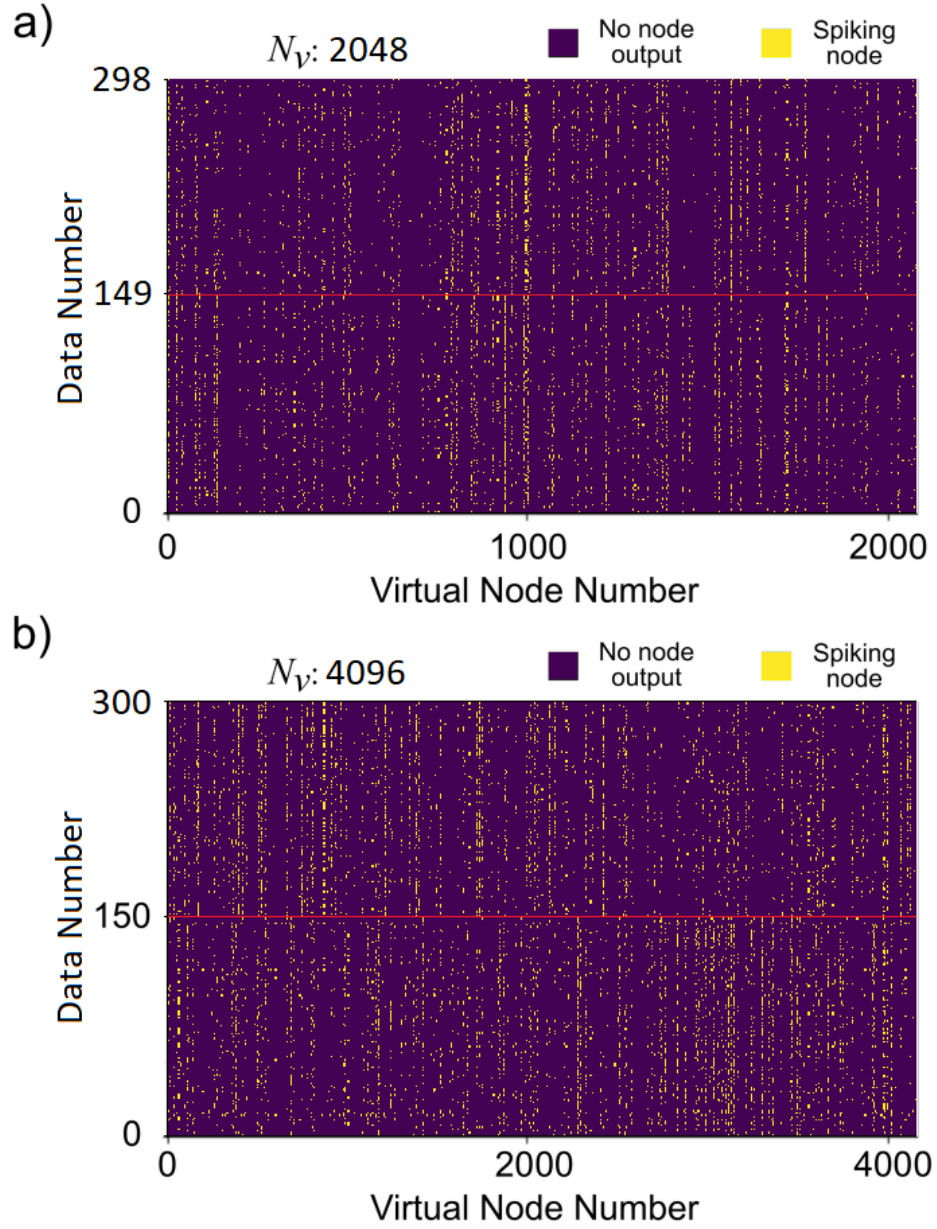


Figure 3.2.2: Temporal maps showing the optical spiking patterns produced by the p-SNN in response to the input dataset using a) 2,048 virtual nodes and b) 4,096 virtual nodes.

set equal, resulting in 298 points. For the network architecture with $N_v = 4096$ (Fig. 3.2.2 (b)), all 300 data points (150 per class) of the MADELON task were used to test the network operation.

Training was done using a randomly selected subset of input datapoints and their corresponding spiking responses. The randomly selected subset contains an equal number of points (N_t) for each class for a total training set size of $2N_t$. The output weights were determined to achieve the best performance on this set with the assumption that the training set is a good representation of the entire task dataset. The input datapoints not used in the training process were subsequently used to test the performance of the MADELON task. The output layer weights calculated during training were applied to the output layer to infer the class of a given test datapoint. The inferred datapoint class was then compared to the true class label, and the performance was measured using classification accuracy, which is the fraction of correctly classified datapoints across all tested datapoints.

3.2.2 Ordinary Least Squares Training Method

First, we again used the ordinary least squares (OLS) training method to test the performance of the photonic SNN in the MADELON task. Here, the output spiking patterns from the VCSEL are treated as a binary sequence of floating-point 1.0s and 0.0s, and the labels are one-hot encoded, and the output weight matrix is found by the product of the Moore-Penrose inverse of the spike vectors and the labels.

Using only 15 training datapoints per class (10 % of the total 300 input datapoints), the OLS training method achieved a very high performance with a peak accuracy of 91% (for a 2,048-node architecture) and 94.5% (for a 4,096-node architecture). Example confusion matrices for $N_v = 2048$ and $N_v = 4096$ trained using this method and $N_t = 15$ are shown in Fig. 3.2.3. As expected, the network architecture with a higher number of nodes (N_v) provided a better overall classification performance.

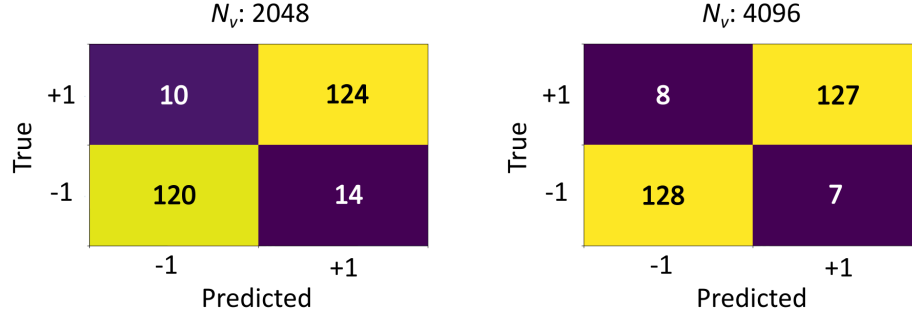


Figure 3.2.3: Example confusion matrices depicting the performance of a single instance of the photonic SNN when trained using the ordinary least squares method. A training set size of 15 ($N_t = 15$) was used to achieved accuracies of 91 % and 94.4 % for architectures of 2,048 and 4,096 nodes, respectively.

3.2.3 Binary Weight Training

While the standard OLS training method delivers a high classification performance during the MADELON task, it does not take full advantage of the discrete binary nature of the fast optical spiking responses in the photonic SNN. For example, it can be observed that in the ideal case of a p-SNN response, spiking patterns would be perfectly consistent between data of the same class, as well as completely distinct from data of alternative classes. In this ideal case, 2^k classes would only require the use of k nodes to be completely separate and identify datapoints. In the ideal case of the MADELON task, only one node per class is required for identification. In the case of experimental data, initially identifying nodes that have more predictive power (i.e., spike only for a particular class) prior to training would allow for the creation of weight matrices that retain high classification accuracy, despite the use and calculation of only a few nodes (a fraction of the total node count).

This work introduces a method of identifying 'useful' nodes using a modified Fisher's chi-squared test statistic[105]. This test statistic is used typically to determine the goodness-of-fit of a distribution to some categorical data, with higher values of the statistic implying less good fit.

The following process is introduced to determine the output weights (the matrix $W_{n,i}$) for this p-SNN:

1. Count the number of spikes that occurred for each class in each node, denoted $S_{n,i}$ (# spikes from data of class i occurring in node n).
2. Calculate the significance score of each node towards each class according to Equation (3.1):

$$Z_{n,i} = S_{n,i}^2 / S_n, \quad (3.1)$$

where $S_n = \sum_i S_{n,i}$ is the total number of spikes occurring in node n across all classes.

3. For each class, choose the top N_n (Node number) scoring nodes and set their weights ($W_{n,i}$) to 1, and all remaining node weights to zero.

The calculated score $Z_{n,i}$ related to the chi-squared test statistic χ^2 through

$$\chi_n^2 = N \sum_i S_{n,i}^2 / S_n - S_n = N \sum_i Z_{n,i} - S_n, \quad (3.2)$$

which would test for goodness-of-fit for the spikes of node n being evenly distributed among the N classes. Thus, higher scores imply that the spikes are not evenly distributed and instead are indicative of the datapoint being of a particular class. The statistic can also be interpreted as a combined score of the number of spikes occurring for a particular class ($S_{n,i}$), and the proportion of spikes occurring in that node only occurring in that class ($S_{n,i}/S_n$).

Applying this algorithm provides the significance scores for all network nodes, with higher scores obtained for nodes that contain not only more spikes, but nodes that have their spikes more concentrated in one specific class. The resulting output weight matrix W contains only nonzero entries on nodes that are strongly indicative of a specific class. The 2D temporal maps in Fig. 3.2.4 reveal the spike trains of a subset of network nodes, for the case when the photonic SNN was operated with a total node count of 2,048. This plot shows the significance scoring system applied to selected nodes. For example, Fig. 3.2.4 shows that node 844 (marked with an amber arrow) elicited an optical spike

for 57 of the total 298 input datapoints used in this experimental run. In this case, the spiking responses from $n = 844$ are evenly distributed across each class (31 spikes for class -1 and 26 for class +1). In this specific case, the significance scores for this node are $Z_{844,-1} = 16.9$ (class -1) and $Z_{844,+1} = 11.9$ (class +1). In parallel, Fig. 3.2.4 shows that Node 885 (marked with a red arrow) fired fewer optical spikes overall. A total of 30 responses were elicited. However, because they were less evenly distributed (and hence more indicative of a particular class), the node scores $Z_{885,-1} = 2.1$ (class -1) and $Z_{885,+1} = 16.1$ (class +1), a class +1 significance score similar to that of node 844. Finally, when the appearance of fast optical spikes is much more biased towards one specific class and is more likely to occur overall, the resulting significance score Z is much higher. For example, in node 853 (marked with a green arrow), 103 spikes occurred, with 88 of them occurring in class -1, yielding scores of $Z_{853,-1} = 75.1$ and $Z_{853,+1} = 2.1$. With this method, the weight $W_{853,-1}$ of node 853 is set to 1, as if a spike was detected in node 853, and it is likely that the datapoint is of class -1. In this example, the weights of the other highlighted nodes ($W_{844,-1}$ & $W_{885,-1}$) are set to zero.

For a data point to be classified, the spike-output vector S must be multiplied by the weight matrix W_i for each class. The class with the highest sum value was the prediction class of the data point. Because the output layer weights are binary, this is equivalent to applying a filter to the optical spike trains and counting the number of remaining spikes. Using this method, the performance can be tuned by selecting the number of nodes (N_n) for training (training node number). Using a fraction of the total number of available nodes can result in equal (if not better) performance, while keeping the weight matrix and number of weight calculations sparse.

Fig. 3.2.5 shows the resulting confusion matrices of the MADELON task when the significance-training algorithm was run with 15 training datapoints ($N_t = 15$). For the p-SNN architecture with 2,048 nodes, a peak accuracy of 92.5 % is obtained, whereas for the 4,096-node architecture a peak accuracy of 95.2 % is achieved. Again, as expected, the architecture with a larger number of nodes provided a higher classification performance. Comparing this performance to the OLS training method, also run at $N_t = 15$,

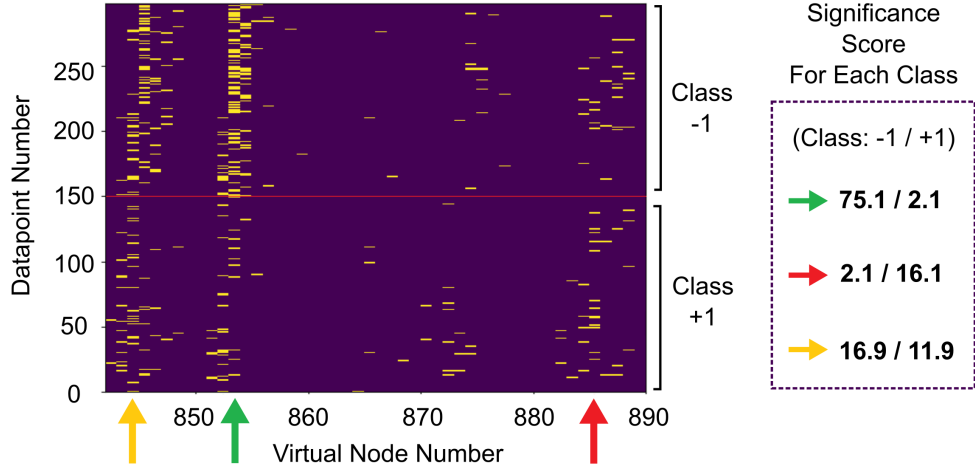


Figure 3.2.4: Section of a 2D temporal map for a network architecture of 2,048 nodes (between nodes 842 and 890). The arrows mark nodes 844 (orange), 853 (green) and 885 (red). In node 853, 88 spikes occurred for class -1, and 15 for class +1, giving it a high significance score of 75.1 (for class -1) and 2.1 (for class +1). Node 844 has significance scores of 16.9 (for class -1) and 11.9 (for class +1). Node 885 has significance scores of 2.1 (for class -1) and 16.1 (for class 1). Node 853 has the highest significance score meaning it would likely be used to perform a classification.

we found that significance training shows a small improvement in peak accuracy in both cases (2,048- and 4,096-node architectures). This improvement in performance shows that despite the same training set size, by considering only nodes significant to specific classes, we can achieve a similar or higher classification performance. Furthermore, it is important to highlight that the training node number used in the significance scoring method was substantially lower than that used in the OLS method. In the novel training method, only the top 20 most significant nodes ($N_n = 20$) were used in both reported p-SNN architecture cases (2,048 and 4,096). Therefore, the novel training method uses only 1% and 0.5% (for 2,048 and 4,096 nodes, respectively) of the total nodes used for OLS training, where every node in the network is trained ($N_n = N_v$). This result indicates that training fewer nodes overall is not only computationally efficient, but does not directly hinder the performance of the system (and in this case, increases classification performance). In Figs. 3.2.6 and 3.2.7, we consider a larger range of training set sizes (N_t), and compare their optimal performance as well as their optimal training node number (N_n).

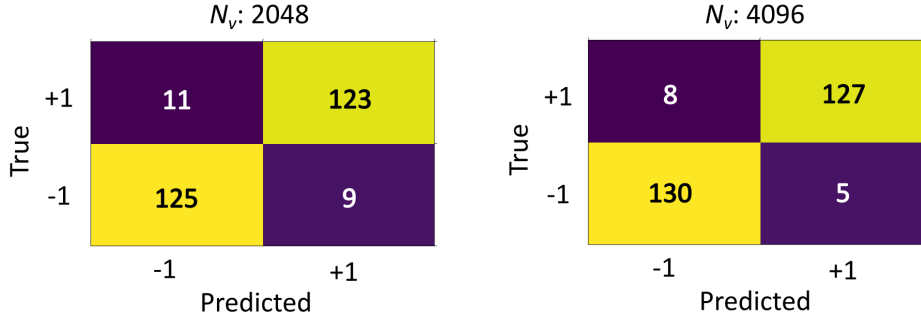


Figure 3.2.5: Example confusion matrices depicting the performance on the Madelon task of a single instance of the p-SNN when trained using the significance score method. A training set size of 15 ($N_t = 15$) was used to achieve accuracies of 92.5 % and 95.2 % for architectures of 2,048 and 4,096 nodes, respectively. Only the 20 most significant nodes ($N_n = 20$) were used to train the p-SNN.

Increasing the number of nodes leads to better performance because of the increased likelihood of a node spiking more strongly for one class than for another. This was caused by the longer-range connections formed in the p-SNN owing to the refractory behaviour of the spiking dynamics in the VCSEL.

Figs. 3.2.6 (a) and 3.2.7 (a) show the peak accuracy achieved in the MADELON classification task as a function of the training set size used for the 2,048 and 4,096 network architectures, respectively. Figs. 3.2.6 (b) and 3.2.7 (b) show the optimal training node number (N_n) used to attain the optimal accuracy.

Both Figs. 3.2.6 and 3.2.7 reveal two key results. First, the p-SNN achieved very high accuracy levels, despite the complexity of the MADELON task and its large number of features. Specifically, a maximum accuracy of 94.4% was achieved for the 2,048-node architecture, and a maximum accuracy of 95.7% was achieved for the 4,096-node architecture. The performance of this p-SNN system on the reduced 300-point data set was comparable to that of machine learning algorithms reported on the full data set as part of the NeurIPS 2003 challenge[104], demonstrating the viability of this platform.

Secondly, the performance of the photonic SNN increases with the training set size (associated with the best-fit training), and the optimal value of N_n may slightly decrease with the training set size. This result was somewhat expected, as training with more data points typically improves the system performance, resulting in a lower

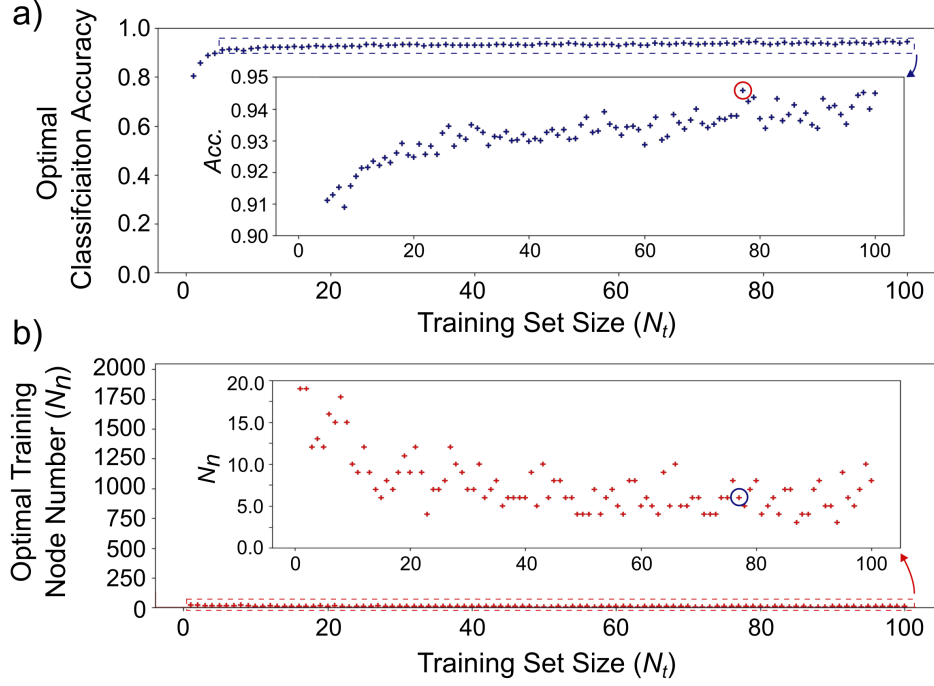


Figure 3.2.6: Photonic SNN performance versus the training set size, using a 2,048-node architecture. a) The optimal classification accuracy (averaged over 10 random selections of the training set) is plotted against training set size (N_t , blue). The inset highlights the consistent high performance over various training set sizes (N_t ranges from 1 to 100). b) The number of training nodes (N_n) used to achieve the corresponding optimal accuracy, plotted against increasing training set size (red). The peak accuracy, 94.4%, occurs when $N_n = 6$ and $N_t = 76$ (circled).

dependence on high node numbers.

Thirdly, Figs. 3.2.6 and 3.2.7 show that the optimal training node number (N_n) remains consistently low while the system retains very high performance, across all training set sizes (N_t) (from 1 to 100). Only small node numbers, as low as < 10 (out of the total node count N_v), must be considered to successfully achieve high performance in the MADELON task. This means that, independent of the training set size, the highly reduced training requirements of the significance scoring method remain a key benefit of the alternative training scheme. Finally, as illustrated in Figs. 3.2.6 and 3.2.7, very small training set sizes, $N_t \leq 10$ ($\leq 3\%$ of the total dataset inputs), are capable of effectively training the photonic SNN system and achieving high accuracy ($> 90\%$) in the MADELON task. Comparing the previous OLS results for a network architecture

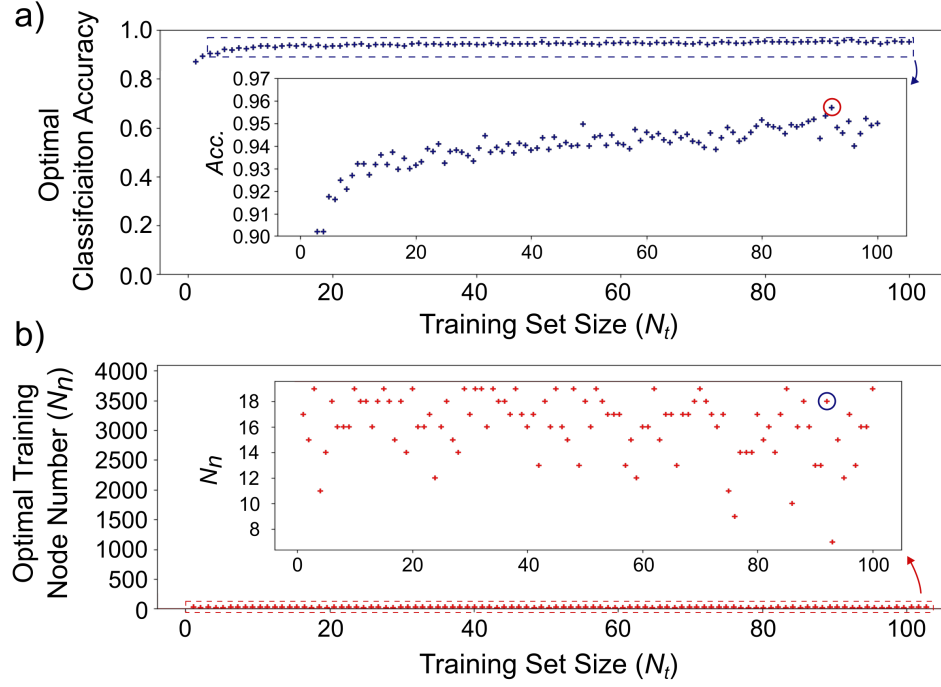


Figure 3.2.7: Photonic SNN performance versus the training set size, using a 4,096-node architecture. a) The optimal classification accuracy (averaged over 10 random selections of the training set) is plotted against training set size (N_t , blue). The inset highlights the consistent high performance over various training set sizes (N_t ranges from 1 to 100). b) The number of training nodes (N_n) used to achieve the corresponding optimal accuracy, plotted against increasing training set size (red). The peak accuracy, 95.7%, occurs when $N_n = 18$ and $N_t = 91$ (circled).

of 4,096 (from 3.2.2) with those in Fig. 3.2.7, we find that at training set sizes < 15 , the significance scoring method achieves similar levels of classification accuracy. Making use of the novel training approach can therefore enable a similar performance to the OLS method, but with the benefit of being more flexible in terms of the number of training datapoints.

The resulting binary weight matrix has several advantages over those produced by typical methods: it is very sparse - all but a few elements are nonzero, making this require very little electronic storage; and as the nonzero elements are only 1 (no floating-point numbers are used), the weights could potentially be directly applied in optical hardware without the need for precision modulation of the signal. Compared to traditional training methods, the algorithm for obtaining this weight matrix is much simpler: typical training methods use costly operations, such as matrix inversions, to calculate floating-point weights, whereas this method simply counts the number of spikes occurring at each node using low-cost computations.

3.3 Time-Series Prediction using a VCSEL p-SNN

Since this p-SNN architecture does not have an optical feedback loop, it lacks internal memory and cannot therefore operate in key tasks, such as time-series prediction, which are of strategic importance in a variety of sectors (e.g. predictive maintenance, weather forecasting, etc.). This section tackles this challenge by demonstrating the successful operation of the p-SNN built with a single VCSEL on a chaotic time-series prediction task. Specifically, we demonstrate experimentally the p-SNN's successful execution of multi-step-ahead prediction of the chaotic Mackey-Glass time-series [106].

This task requires that the system produces an output corresponding to a real number (the time-series value), and that the system has knowledge of previous temporal data inputs (memory). This requires going beyond the simpler operation in classification tasks for which memory is not required and whose output is limited to an element of a finite set (equal to the number of datapoint classes). Here, we demonstrate the successful performance of the p-SNN overcoming these fundamental challenges caused

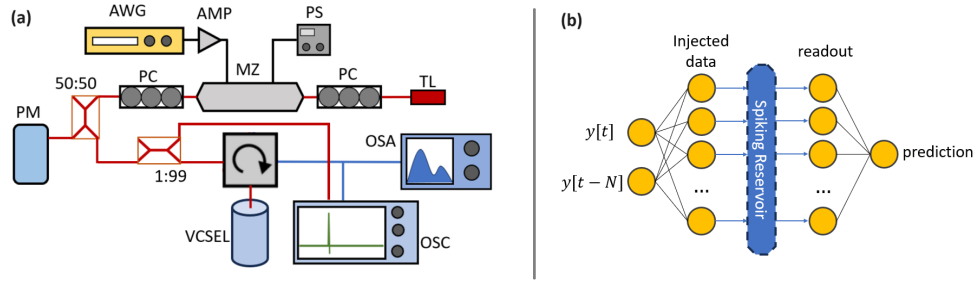


Figure 3.3.1: Experimental Layout and Logical Architecture of the VCSEL-based photonic Spiking Neural Network. (a) Diagram of the experimental photonic spiking neural network. A tuneable laser (TL) provides light which is intensity modulated by a Mach Zehnder (MZ) modulator. The input to the MZ modulator is given by a DC bias (generated by a Power Source, PS) and an arbitrary waveform generator (AWG) followed by an RF amplifier (AMP), encoding the data to be processed into the time varying intensity of the light. Polarisation controllers (PC) are used to match the light polarisation of the input signal to that of the MZ modulator and the VCSEL. The input signal is monitored with a Power Meter (PM) and oscilloscope (OSC). The modulated light signal is then injected into the VCSEL where it can trigger ultrafast optical spikes. The VCSEL's output is read by an optical spectrum analyser (OSA) and a photodiode and oscilloscope (OSC). (b) Schematic diagram of the architecture of the VCSEL-based p-SNN [66, 67]

by the representation of real numbers with discrete neural-like spikes, and the lack of (memory enhancing) feedback network connections. We achieve this by applying a memory-inducing technique in the pre-processing of the input data that has recently been shown to be an effective means of improving the performance of reservoir computers on tasks requiring memory [107].

3.3.1 Experimental Results on the time-series Prediction Task

time-series prediction is the task of predicting a future value in a sequence of numbers in a time-series, given the preceding values. This is an interesting task for different key applications, such as predictive maintenance and weather forecasting. However, for sequences (time-series) defined by complex or chaotic dynamics, multi-step-ahead predictions can be a very challenging task, as subsequent values are not strongly correlated to previous historic values in the sequence.

Time-series prediction has been previously investigated in traditional (non-spiking)

photonic reservoir computing systems, especially in photonic implementations based upon time delay reservoirs (TDR) using semiconductor lasers, and also including VCSELs as core nonlinear elements (see for example [44, 46] and references therein). In such systems, there is an explicit optical delayed-feedback loop (typically implemented by an optical fibre link), forming recurrent connections and therefore allowing data injected in previous steps to continue circulating in the system; hence creating internal memory and influencing the output and subsequent predictions. This memory effect allows photonic RCs to combine information from several data points in the input time-series to make accurate future-value predictions [108, 109].

However, since the VCSEL-based p-SNNs do not include any intrinsic mechanism to create recurrent connections, the internal system memory inherent to photonic TDRs is not present in the VCSEL p-SNN, limiting their use for time-series prediction functionalities. We tackle this key challenge proposing an alternative data encoding approach allowing to create memory effects in the p-SNN. This is realised during the input data stage, by adding to the original data input a delayed copy of the time-series. In this situation, at a given time-instant the input information contains two data input points, namely the present and a delayed value [107], using which the system is able to make multi-step-ahead predictions of a time-series. With this simple, yet very powerful approach, information about the long-range dynamics of the time-series to be predicted is made available to the p-SNN, thus effectively creating memory in the system. Notably, this is done without the need to add external fixed optical delay feedback lines, which increases complexity, require precise temporal control and can limit the total processing rate of the system. Additionally, the number of nodes can be changed at will due to the feedforward p-SNN architecture of the network. This permits the user to easily set diverse networks with higher or lower node counts depending on the complexity of the task to be processed.

To investigate the operation of the VCSEL p-SNN for time-series prediction, we focus on demonstrating the system’s ability to perform multi-step-ahead predictions of the Mackey-Glass chaotic time-series, defined by the delay differential equation in Equation 3.3 [106]:

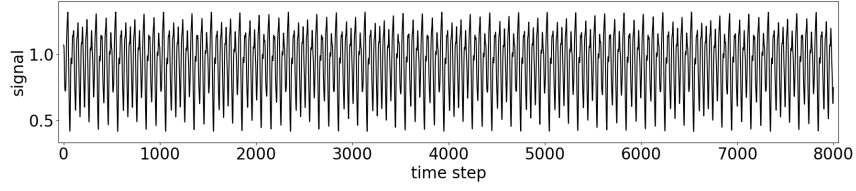


Figure 3.3.2: The Mackey-Glass time-series. Time-trace showing the 8000 data points of the chaotic Mackey-Glass time-series used for the experimental investigation of the next-step ahead prediction task in the VCSEL-based p-SNN.

$$\frac{dy}{dt} = \frac{\beta_0 y(t - \tau)}{1 + y(t - \tau)^n} - \gamma y(t), \quad (3.3)$$

in which the values β_0 , γ and n are parameters that take standard values $\beta_0 = 0.2$, $\gamma = 0.1$ and $n = 10$, and the delay $\tau = 15$.

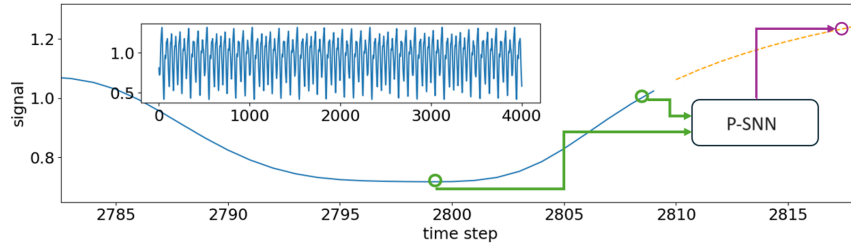


Figure 3.3.3: Operation of the p-SNN in the time-series prediction task. The time-series defined by the Mackey-Glass equations is shown in blue and dashed orange. The task is to use the values before a particular time step (in this case time-step 2809) to predict the value several steps ahead. To do this, the value at the time step and the value 10 steps preceding (both marked with green circles) are combined and processed in the p-SNN to find a prediction (represented with the purple circle). The inset shows half of the full time-series.

For the experimental demonstrations of this work we used a data set of 8000 inputs and targets to demonstrate the time-series prediction functionality of the VCSEL p-SNN. Figure 3.3.2 shows the 8000 time-series point sequence used as data input. To encode the inputs of the Mackey-Glass time-series for their optical injection into the VCSEL p-SNN, each given data point was combined into a two-component vector with the data point delayed by 10 steps (as shown graphically in Figure 3.3.3). These resulting vectors were then multiplied by a randomly chosen input weight 'mask' to

give the input values for each of the virtual nodes forming the p-SNN. The input mask was thus a matrix with a dimension of $2 \times N_v$, where N_v describes the total number of nodes in the network, with entries drawn uniformly at random from the interval $[0,1)$. The resulting masked input was then normalised so that all values were in the range $[0,1)$ and were concatenated to generate the RF signal with the input data to be encoded optically (with the Mach-Zehnder Modulator in the setup) for injection into the 1300nm-VCSEL at the core of the p-SNN. The input delay of 10 steps was chosen to enable multi-step-ahead predictions resulting in predictions horizons of about 10-time units of the Mackey-Glass system.

Figs. 3.3.4(a) and 3.3.4(b) show typical input and output optical signals from the VCSEL p-SNN after their analysis with the high-speed photodetector and real time oscilloscope in the setup. In particular, fig. 3.3.4(a) depicts the temporal varying nature of the optical input signal entering the p-SNN with the encoded Mackey-Glass data points. Fig. 3.3.4(b) shows in turn the high-speed optical spike firing signal obtained at the output of the p-SNN in response to the input shown in fig. 3.3.4(a). For the results in Figs. 3.3.4(a) and 3.3.4(b), the p-SNN was set with a node time of 250 ps and a total network node count of $N_v = 1024$; yielding a total processing time of 256 ns per data-point. Figs. 3.3.4(a) and 3.3.4(b) depict the input and output signals of the p-SNN when two consecutive data input points of the Mackey-Glass time-series are injected into the system.

The Mackey-Glass chaotic time-series prediction task was run on the VCSEL p-SNN using three total node counts, namely $N_v = 256, 512$ and 1024 , which for a set node time of 250 ps gave processing times of 64, 128 and 256 ns per data point respectively. In all cases of analysis, the 1300 nm VCSEL at the core of the p-SNN was biased with a driving current of 3.51 mA at a temperature of 293 K, and the detuning of the injection frequency from the peak frequency of the subsidiary (orthogonally-polarised) mode of the VCSEL was -3.6 GHz. The average injection power in each case was $127 \mu W$, $123 \mu W$ and $131 \mu W$ for the p-SNN architecture set with 256, 512 and 1024 nodes, respectively.

To train the p-SNN to perform the chaotic Mackey-Glass time-series prediction

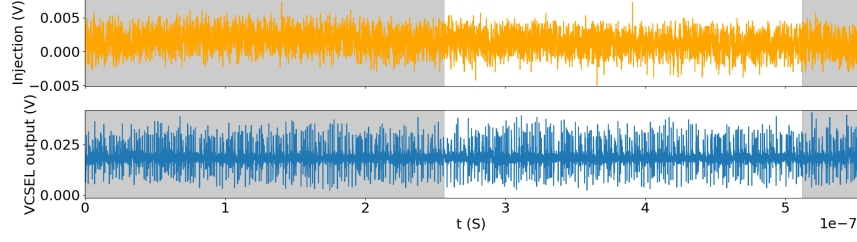


Figure 3.3.4: Input and output signals. Photodetector recording of the intensity of light injected and resulting intensity output from the VCSEL p-SNN showing the triggered ultrafast (sub-nanosecond long) optical spikes. The p-SNN was set here with a total node count of $N_v = 1024$; thus, giving a total processing time per point of 256 ns. The plots show the trace for two data points of the Mackey-Glass time-series lasting from 0 ns to 256 ns (shaded region), and from 256 ns to 512 ns (unshaded region) and the achieved optical spiking outputs.

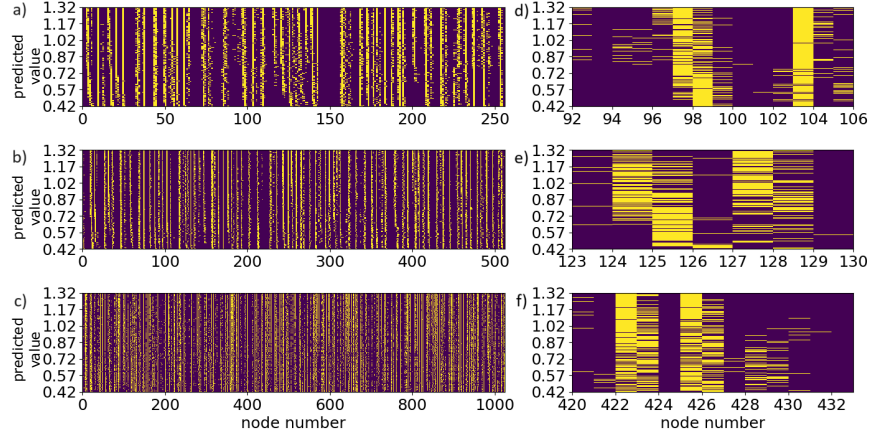


Figure 3.3.5: Spiking Patterns generated by the VCSEL. Temporal maps depicting the optical spike patterns delivered by the VCSEL p-SNN in response to the input data values of the Mackey-Glass time-series, sorted by the value to be predicted. Three different network architectures with total node count (N_v) of $N_v = 256$ (a,d), $N_v = 512$ (b,e) and $N_v = 1024$ (c,f) are investigated. The panels (d,e,f) provide zoomed in regions of the temporal maps that demonstrate differences in the optical spiking patterns achieved in different network nodes for different values of the Mackey-Glass time-series. A small difference in the spike pattern (i.e. a node begins/stops spiking) corresponds with a small difference in the target value to be predicted.

task, Ordinary Least squares fitting was used. The optical spike vectors obtained from the p-SNN for all input data points used for training are collected into a state matrix. Then, the state matrix and vector of target time-series values are used to find the output layer weight matrix. After the training process, during the testing phase, the binary vectors with the p-SNN's output spiking response obtained for subsequent input data points, are multiplied by the calculated output layer weight matrix, to obtain the predicted values of the Mackey-Glass chaotic time-series. This process permits the transformation of the binary (spiking/non-spiking) temporal output of the p-SNN into a real number providing the next-ahead predicted value. The spike pattern vectors for the entire dataset used in this work is shown in Fig. 3.3.5, sorted according to the value that is to be predicted. Different nodes are active (i.e. contain spikes) for different ranges of the value, which is learned by the least-squares regression fitting.

The performance of the p-SNN on multi-step-ahead prediction of the Mackey-Glass chaotic time-series for prediction horizons of 1 to 15 is shown in Fig. 3.3.6. The accuracy of the p-SNN has been measured using the Normalised Mean Squared Error (NMSE) which is calculated using the relation provided in equation 3.4. There, the true values are denoted y and the prediction is denoted $\hat{y}$. Fig. 3.3.6 shows that despite the lack of intrinsic hardware memory of the p-SNN and the limited temporal information provided to it, very good prediction performance is achieved over the entire prediction range. A local minimum in the prediction error is evident around 10-steps-ahead, where a low NMSE value equal to only 0.051 is achieved. This is because the input delay of 10 used in this setup, is optimised to 10-step-ahead predictions [107]. The lowest errors are achieved for one-step-ahead predictions (with NMSE reducing to only 0.021), as the time-series values are highly correlated over this time range. We selected the NMSE metric as this has been widely used to analyse accuracy in traditional photonic RC systems (including those recently reported based upon VCSELs), in chaotic time-series prediction tasks. Hence, we also use it here to allow for a better comparison between the performance of the VCSEL p-SNN of this work with other photonic approaches.

$$NMSE = \frac{\sum_{i=0}^N (y_i - \hat{y}_i)^2}{N\sigma_y^2} \quad (3.4)$$

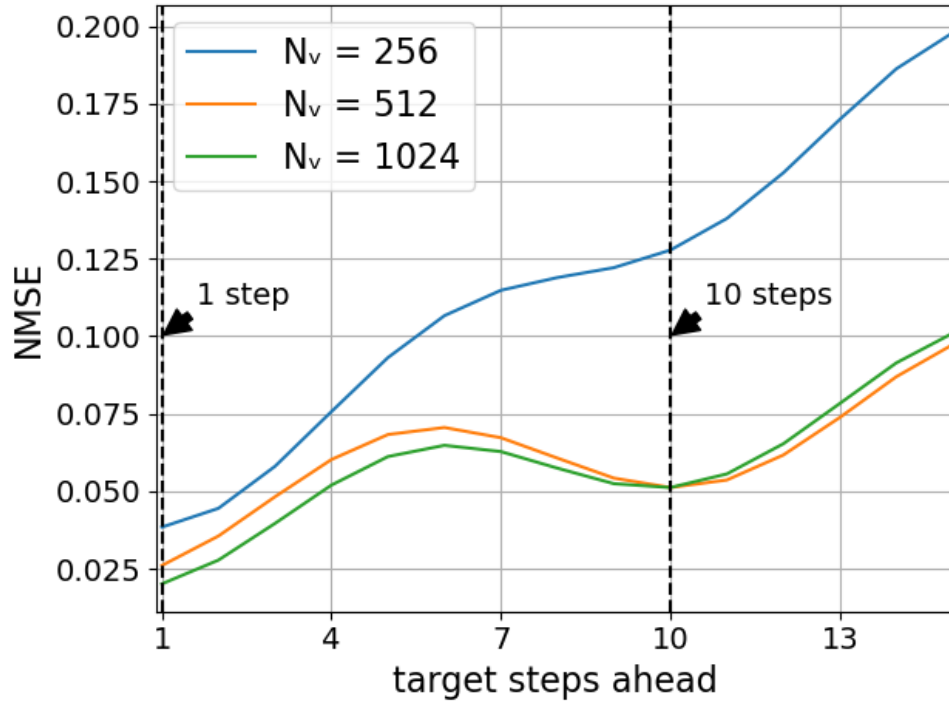


Figure 3.3.6: Multi-step-ahead prediction performance. Performance of the VCSEL p-SNN using input pairs comprised of the current value of the Mackey-Glass time-series and the value delayed by 10-steps for each node number (N_v) used. The x-axis is the number of steps ahead of the target and the y-axis gives the accuracy of the predictions in terms of the Normalised Mean Squared Error (NMSE). Training and testing parameters: training set size is 6000, testing set size is 1500 and the regularisation parameter is 0.03.

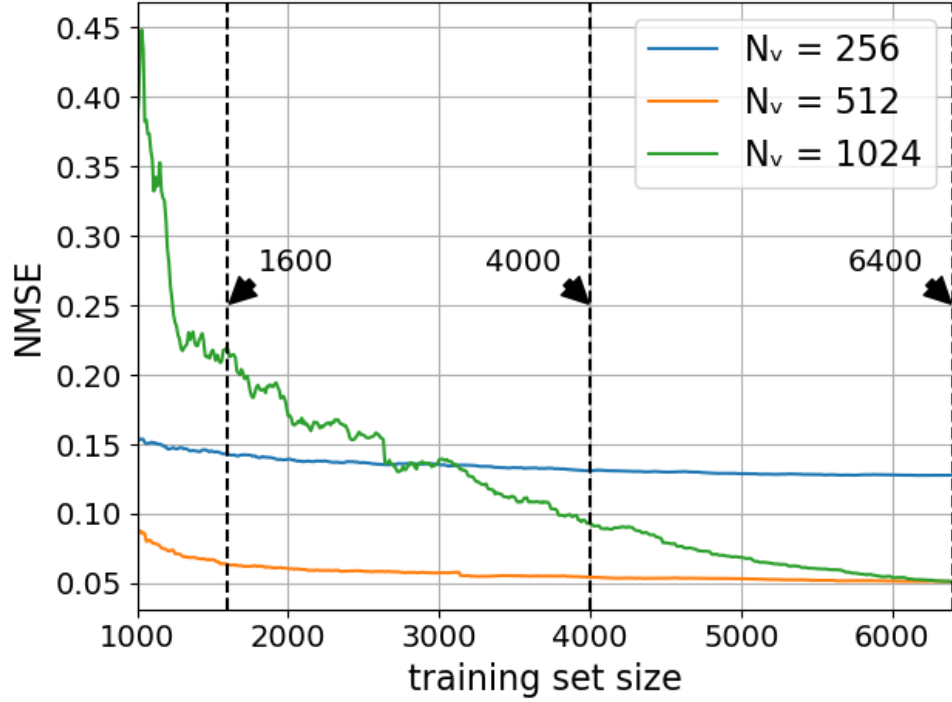


Figure 3.3.7: Performance against Training set size. Performance of the VCSEL p-SNN in the 10-step-ahead prediction of the Mackey-Glass time-series as a function of the training data set size, for each node number (N_v) used. The prediction error is the Normalised Mean Squared error (NMSE). Training and testing parameters: testing set size is 1500 and the regularisation parameter is 0.03.

Fig. 3.3.7 shows the influence of the training data set size for 10-step-ahead predictions. Here we find very good performance for training sets as small as 1600 datapoints (20% of total data points applied to the system) using the p-SNN with a 512-node network architecture. Fig. 3.3.7 also shows that for a larger network node count of 1024 nodes, larger training sets are needed to avoid overfitting.

To highlight the system operation, detailed results are provided in this work for three different training data sizes, namely 1600, 4000 and 6400 points (20%, 50% and 80% of total), and for the three different network architectures (with 256, 512 and 1024 nodes) investigated. Fig. 3.3.8 plots as an example of the VCSEL p-SNN performance in the Mackey-Glass time-series prediction task for both 10 steps-ahead (fig. 3.3.8(a)) and 1 step-ahead (fig. 3.3.8(b)) predictions. For the results in figs. 3.3.8(a) and 3.3.8(b), a dataset formed by the first time-ordered 4000 points (50% of the total 8000

points applied) in the time-series is used for the training phase, and the last 4000 points are used during inference. For the specific cases investigated in Fig. 3.3.8 the p-SNN was configured with a total node count of 512 nodes, which at 250 ps/node, yields a total processing time of 128 ns per data point. The blue and orange traces in Figs. 3.3.8(a) and 3.3.8(b) plot respectively the original time-series and the predicted values delivered by the p-SNN. Figs. 3.3.8(a) and 3.3.8(b) show that very good performance is obtained in both the 10 and 1 steps-ahead prediction tasks with the p-SNN. This is further illustrated in the insets in Figs. 3.3.8(a) and 3.3.8(b) showing in detail different smaller temporal sequences, highlighting the high accuracy performance in the time-series prediction task delivered by the p-SNN, even when a very small dataset size (50% of total points) is utilised for training and with all data points fed in a timely-ordered manner. The accuracy of the p-SNN has also been measured using the correlation between true and predicted values. To that end we calculated the Pearson correlation coefficient using equation 3.5. There, the true values are denoted y and the prediction is denoted $\hat{y}$. This measures how well the predictions depend linearly on the true value (a Pearson coefficient of 1.0 implies perfect linearity between the two variables). The plots at the right side of Figs. 8(a) and 8(b) provide the correlation between true and predicted values for the two examples considered, further highlighting the good performance of the system in predicting different step-ahead values of the Mackey-Glass time-series.

$$\rho = \frac{\text{cov}(y, \hat{y})}{\sigma_y \sigma_{\hat{y}}} \quad (3.5)$$

Table 3.1 shows the performance of the p-SNN in the 10 steps-ahead value prediction of the chaotic Mackey-Glass time-series, measured from the calculated NMSE and correlation coefficients for the 3 specific cases of training dataset sizes of 20%, 50% and 80% (as percentage of the 8000 points used) and for network architectures with $N_v = 256$, investigated in this work. The results in Table 3.1 highlights that for all cases of analysis high-accuracy is found in the 10 steps-ahead prediction of the Mackey-Glass time-series, as per the small obtained values of NMSE (as low as $\text{NMSE} = 0.051$, when

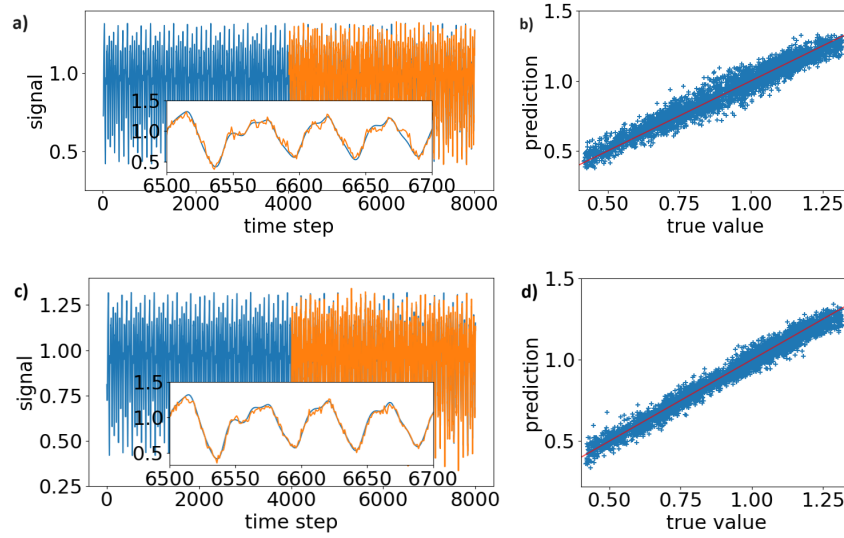


Figure 3.3.8: Examples of time-series prediction. Performance of the p-SNN in the prediction of step-ahead values of the Mackey-Glass time-series using a network with 512 virtual nodes and applying a training dataset using 4000 points (50% of total). (a,b) Predicting 10 steps ahead. (c,d) Predicting the next step ahead. (a,c): Predicted (orange) and true (blue) time-series values, with the first 4000 points shown in blue providing the training dataset. The insets in plots (a) and (c) show a zoomed in region between data points 6500 and 6700 of the Mackey-Glass time-series showcasing in more detail the good system performance in this prediction task. (b,d): Correlation between true and predicted values (in red is shown the line where prediction = truth).

Node Count	Training set %	NMSE	Correlation
256	20	0.148	0.91
	50	0.134	0.92
	80	0.131	0.92
512	20	0.066	0.97
	50	0.055	0.97
	80	0.051	0.97
1024	20	0.213	0.87
	50	0.087	0.95
	80	0.051	0.97

Table 3.1: Error and correlation against training set size. Experimentally obtained values of NMSE (Normalised Mean Squared Error) and correlation coefficient for 10-steps ahead predictions for three different configured network node architectures in the p-SNN and different training set sizes (as a percentage of 8000).

$N_v = 512$) and high levels of correlation (up to 0.97 is achieved in multiple cases of analysis) between true and prediction values. While the performance of this p-SNN is comparable to that of typical photonic reservoir computing platforms [46, 54], the p-SNN again has the ability to change the number of nodes (and thus processing speed and performance) at will.

3.4 Summary

The works presented in this chapter have thus demonstrated the use of optically injected VCSELs as the active element at the centre of a neural network. The spiking response of the VCSEL to modulation of the input injection signal effectively implements the activation function of the neural network, through a combination of the leaky-integrate-and-fire and refractory dynamics. This neural network was then applied to tasks such as data classification and time-series prediction, and a novel training algorithm was introduced, taking advantage of the discrete output of the spiking neuron.

Chapter 4

Resonant Tunnelling Diode Neurons

This chapter introduces a device known as the Resonant Tunnelling Diode (RTD) and its use as an artificial spiking neuron for use in neuromorphic processing technologies. The RTD is a type of semiconductor device centred around a double-barrier quantum well within its band structure, notable for showing an effective "negative differential resistance". In these devices, the quantum tunnelling of electrons is directly observable at room temperatures, showing up as highly nonlinear, non-classical conductance characteristics. Circuits constructed around these devices can thus exhibit ultrafast dynamics, being dependent only on the speed of electron tunnelling, ideal for integration alongside purely photonic systems.

Here, I will describe the semiconductor structure and mechanism underlying the working of RTDs, describe the possible ways to couple these devices to photonics (through photodetector and lasers), and set out a numerical model of an optoelectronic RTD to allow simulation of these devices. The operation of RTDs as artificial spiking neurons, showing the typical features such as spike firing, threshold-and-fire and refractoriness will be demonstrated, as well as rich dynamics that can be used for neuromorphic processing.

All the numerical simulation and experimental work on these RTDs and the charac-

terisation for neuromorphic processing systems was carried out by me at the University of Strathclyde, while the RTD devices were fabricated by research collaborators at external institutions, namely the University of Glasgow, UK and the Eindhoven University of Technology (TU/e), The Netherlands.

4.1 The Physics of Resonant Tunnelling Diodes

A Resonant Tunnelling Diode (RTD) consists of a quantum well sandwiched between two high-energy barriers, formed by layers of different semiconductors. According to the wave-like treatment of electrons, the well in the centre of the RTD has a set of discrete trapped energy levels, and electrons incident on the barrier structure are able to tunnel into and out of these levels with probability according to their energy. An important result is that for the symmetric double barrier well, electrons with the right energy can tunnel through the entire structure with 100% probability - a phenomenon known as resonant tunnelling[110] (hence the name of the device).

The conduction band structure of a typical RTD is shown in figure 4.1.1. The two barriers are formed by a semiconductor material with a wide bandgap, increasing the energy of the conduction band, while the well is formed by a low bandgap material. The widths of each section is very thin, the barriers are typically $20 \text{ \AA}$ and the well $60 \text{ \AA}$ [111], in order to maintain the coherence of the electron wavefunction across the structure, even at room temperature.

The conduction across this structure is then determined by the ability of electrons to tunnel through the well. At low applied voltage, the incoming electrons have an energy lower than that of the lowest bound state in the well, and the tunnelling probability is suppressed. As the voltage and thus incident electron energy increases, the energy difference decreases, until the incident electron energy is that of the trapped level, and tunnelling is resonant, allowing high current to flow through the diode.

Beyond this however, the energy difference between the conducting electrons and the well states increases again, thus decreasing tunnelling probability and the resulting current. This non-ohmic decrease in current with increase in voltage is termed

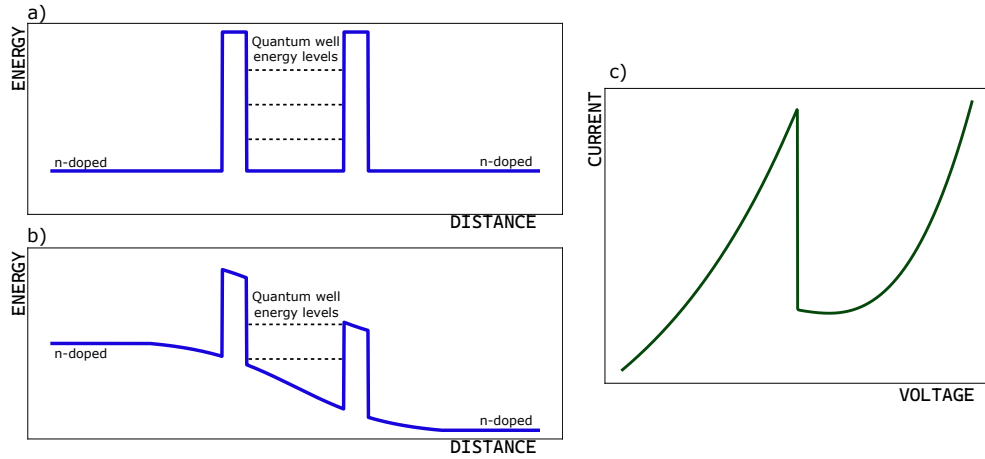


Figure 4.1.1: The conduction band of an RTD under (a) zero voltage, (b) external reverse voltage bias. The two barriers are formed by wide-bandgap semiconductors, and give rise to the well in the centre that holds discrete electron energy levels. As the external voltage is increased, the energy of the band outside the well becomes closer to the level of one of the trapped states, facilitating resonant tunnelling. (c) The shape of the resulting conductance curve. The decrease in current with voltage in the middle, resulting from the drop in tunnelling probability when the applied voltage is increased above the trapped state, is known as negative differential resistance (NDR).

”negative differential resistance” (NDR), and is the key feature of RTDs that allows their operation as ultrafast oscillators, and even spiking neurons. Depending on the size of the well and the materials used, an RTD can display multiple voltage ranges of NDR, though typically devices only have one. At much higher voltages, the electrons have energy higher than the barriers and so are conducted as if through a simple PN junction.

RTDs can also be made optically active, by incorporating an ”optical window” into the structure of the diode [112–114], allowing light to reach the photosensitive layer. In these devices, light incident on the diode causes the emission of carriers into the conduction band, which flow as current through the circuit. Additionally, accumulation of carriers (either electrons or holes) on one side of the well can enhance the potential drop across the well structure, further changing the conduction characteristics.

The following sections will describe the RTD-based circuit that will be used for neuromorphic processing and a mathematical model describing its voltage and current dynamics, alongside photonic elements such as a laser and photodetector.

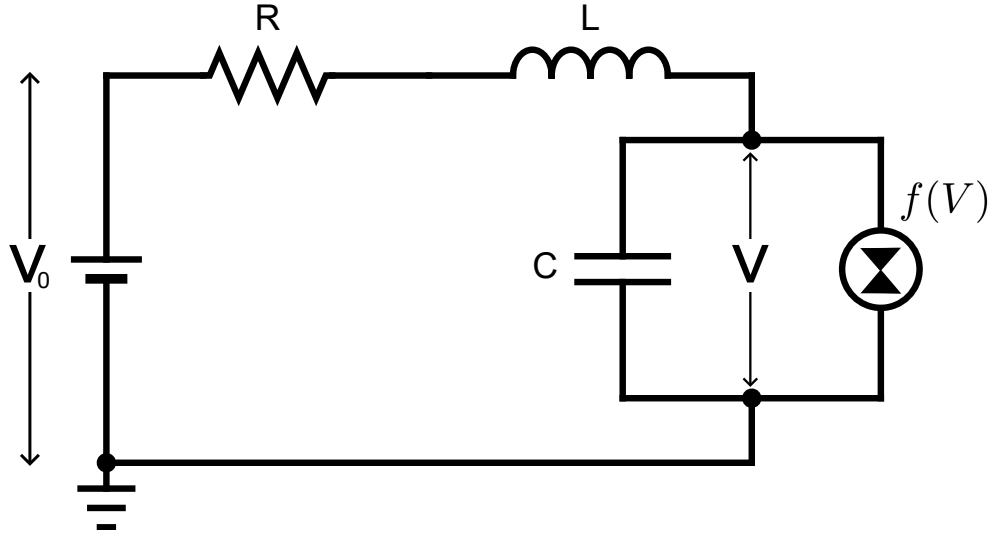


Figure 4.1.2: Lumped Circuit Model used to describe the RTD. The external cables and power supply each contribute a parasitic resistance, capacitance and inductance, labelled by R , L and C . The nonlinear conductance characteristic of the RTD is represented as a voltage-dependent resistor (where the resistance is defined by a function $f(V)$). Analysis of this circuit using classical laws gives a good model of the dynamics of a real RTD.

4.1.1 Resonant Tunnelling Diode Circuits

Figure 4.1.2 shows a schematic of the simplified circuit that models a single RTD connected to a variable voltage supply, in which the RTD is modelled as a resistor with a voltage dependent conductance, given by a function $f(V)$. The remainder of the circuit (e.g. cables, power supply and connecting components) is represented by a combined inductance, capacitance and resistance, resulting in a circuit similar to a radio oscillator.

Analysis using Kirchhoff's laws gives the following description of the dynamics of the circuit:

$$C\dot{V} = I - f(V) \quad (4.1)$$

$$L\dot{I} = V_0(t) - V - RI, \quad (4.2)$$

in which V is the voltage drop across the RTD, I is the current flowing in the

circuit, and V_0 is the voltage supplied by the power supply. These coupled equations describe an oscillator in the (V, I) phase space, with the shape of the orbits defined by the shape of $f(V)$

This two-dimensional first-order system can also be reduced to a one-dimensional, second order system known as the Liénard oscillator[115, 116], by differentiating equation 4.1 and substituting, yielding:

$$\ddot{V} + p(V)\dot{V} + q(V) = 0, \quad (4.3)$$

where $p(V) = \frac{R}{L} + \frac{1}{C}f'(V)$ and $q(V) = \frac{1}{LC}(V - V_0 - Rf(V))$. The Liénard oscillator is particularly notable for modelling spiking neurons due to the range of behaviours exhibited by the system including relaxation oscillations and limit cycles, excitability, and bursting, each of which is observed in biological neurons.

4.1.2 Current-Voltage characteristics of RTDs

The accuracy of the lumped circuit model depends on the form of the current-voltage (IV) curve given by $f(V)$. This describes the amount of current that flows through the double-barrier structure given the voltage across the device.

The simplest function that shows an NDR region is a cubic polynomial e.g. $f(V) = aV^3 - bV$. This choice of $f(V)$ corresponds to the Van de Pol oscillator[115, 117], used to describe oscillations with nonlinear damping found in vacuum-tube valve based circuits.

A cubic curve however does not fit well the observed shape of the RTD IV curve, which has a much steeper drop in the NDR. A physics-based form of $f(V)$ was derived by Schulman[118], considering the electron transport through the double-barrier structure:

$$f(V) = A \ln \left[\frac{1 + e^{(q/k_B T)(b-c+n_1 V)}}{1 + e^{(q/k_B T)(b-c-n_1 V)}} \right] \times \left[\frac{\pi}{2} + \tan^{-1} \left(\frac{c - n_1 V}{d} \right) \right] + H \left[e^{(q/k_B T)n_2 V} - 1 \right] \quad (4.4)$$

In this model, q, k and T are the electron charge, Boltzmann constant and tem-

Parameter	Value
A	$-5.5 \times 10^{-5} A$
b	$0.033 V$
c	$0.113 V$
d	$-2.8 \times 10^{-6} V$
n_1	0.185
n_2	0.045
H	$18 \times 10^{-5} A$

Table 4.1: Parameters used in the definition of the IV curve.

perature respectively, A and H are currents, b and c are voltages, and n_1 and n_2 are dimensionless constants. These parameters can be tuned to match experimental data, however they do not explicitly correspond to physical dimensions of the RTD, or to fabrication parameters.

The IV curve defined by this function has two turning points: a local maximum, referred to as the "peak", followed by a local minimum, known as the "valley", between which lies the NDR voltage range. The voltage and current values of these points are useful for characterising the RTD for various uses: the ratio of the currents of the peak and valley is often reported when the RTD is used as an oscillator or amplifier[112, 119], while a larger difference between the voltages of the peak and valley will be useful when the RTD is used as a spiking neuron.

Table 4.1 shows the values of the parameters of $f(V)$ used for modelling of RTD neurons in this work, fitted to experimental results in a previous work[120], and the corresponding IV curve is shown in figure 4.1.3 (a). The NDR region is shaded, ranging from $0.61V$ to $0.72V$. The effect of varying parameters A , H and c is shown in figure 4.1.3 (b-d)

4.1.3 RTD Circuit Dynamics

Equations 4.1 and 4.2 define the flow of trajectories in the V, I phase space. The nullclines of the dynamical equations are the sets of points where each equation is zero, and the time derivatives of voltage or current is zero. These occur on the curve $I = f(v)$ (for the $\dot{V}$ nullcline) and the line $I = (V_0 - V)/R$ (for the $\dot{I}$ nullcline). Where

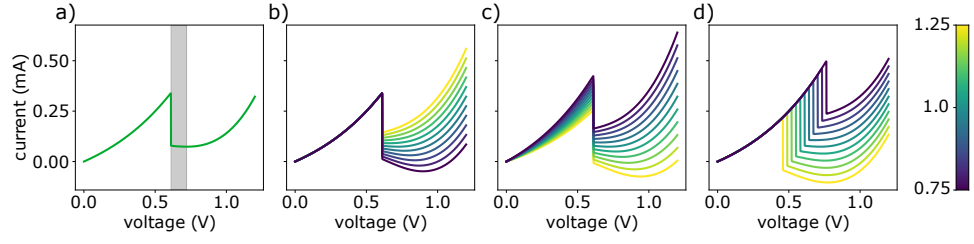


Figure 4.1.3: RTD IV curves. a) The IV curve defined by parameter values in table 4.1. The region of NDR runs from $0.61V$ to $0.72V$, shaded. b-d) Samples of RTD IV curves for variations of some of the parameters in equation 4.4 around the default values used in this work. The panels show the variance of b) A , c) H , and d) c . Each parameter is varied between 75% and 125% of their defaults, while all other parameters are not changed.

these cross, the system has a dynamical fixed point, with voltage (V^*) and current (I^*) given by $f(V^*) = (V_0 - V^*)/R$ and $I^* = f(V^*)$. Typically, there is only one fixed point, however for low values of R , the current nullcline is sufficiently shallow that it can cross the IV curve in multiple points (up to three points). Figure 4.1.4 plot the nullclines and trajectory flows for several values of the bias voltage V_0 , showing the flow of the circuit around the IV curve and convergence onto the stable fixed point.

The dynamics around the fixed point at the intersection of the nullclines can be analysed by linearising the nonlinear system, considering the Jacobian of equations 4.1 and 4.2:

$$J(V, I) = \begin{bmatrix} -\frac{f'(V)}{C} & \frac{1}{C} \\ -\frac{1}{L} & -\frac{R}{L} \end{bmatrix} \quad (4.5)$$

This can be reduced to dynamics of the dimensionless quantity $X = \dot{V}$ yielding the one-dimensional second-order equation:

$$LC\ddot{X} + (RC + Lf'(V^*))\dot{X} + (1 + Rf'(V^*))X = 0 \quad (4.6)$$

If a fixed point has coordinates (V^*, I^*) , the point is stable if the trace of $J(V^*, I^*)$ is negative, and the determinant positive, and unstable otherwise, which occurs when $f'(V^*) > -RC/L$ and $f'(V^*) > -1/R$. This is trivially true outside the NDR, where

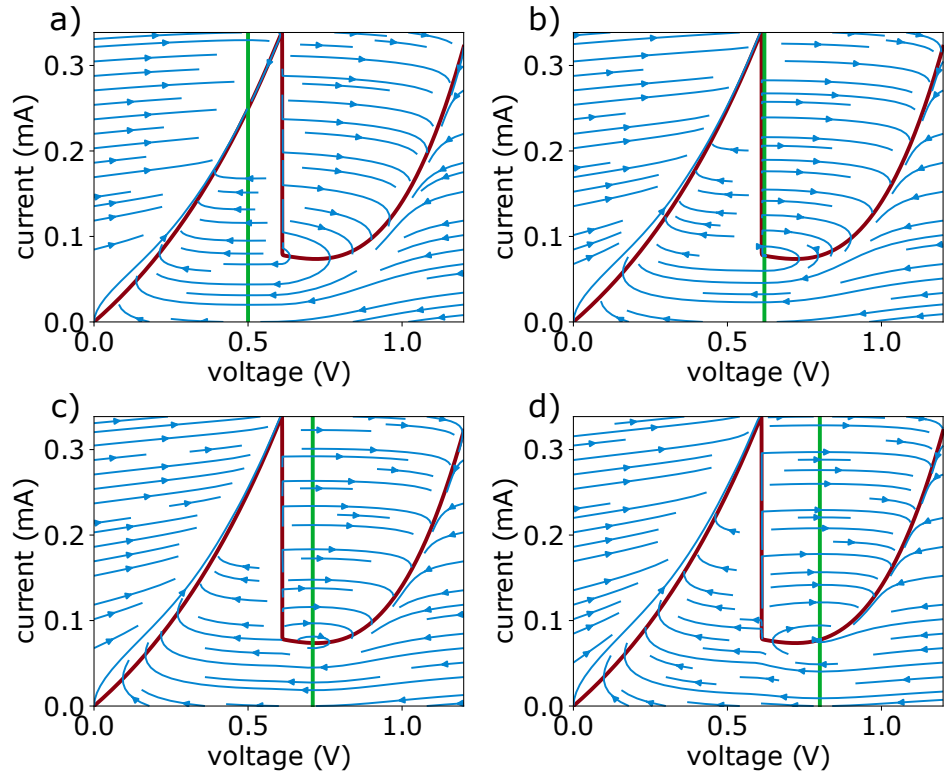


Figure 4.1.4: Nullclines (voltage in red, current in green) of the RTD circuit system when biased at different voltages, and flow of the trajectories. The points at which the nullclines cross is a fixed point of the system. The RTD was biased at a) $0.5V$, b) $0.61V$, c) $0.71V$ and d) $0.8V$. Note that in b), the fixed point is inside the NDR, and is unstable. In all other cases, the fixed point is stable, and the flows converge on that point.

$f'(V)$ is always positive, so all unstable fixed points must occur inside the NDR.

Equation 4.6, which describes the dynamics of the RTD circuit within a neighbourhood of a fixed point, describes a damped harmonic oscillator with undamped frequency ω_0 and damping coefficient ζ :

$$\omega_0 = \sqrt{\frac{1 + Rf'(V^*)}{LC}}, \quad (4.7)$$

$$\zeta = \frac{RC + Lf'(V^*)}{2\sqrt{LC + RLCf'(V^*)}}, \quad (4.8)$$

which has solutions of the form $X(t) = X_0 e^{i(\omega_0 \sqrt{1-\zeta^2} + i\zeta\omega_0)t}$. The stability conditions of the fixed point can then be understood as the requirements that the frequency ω_0 is real and damping ζ is positive, as otherwise $X(t)$ would diverge. This also differentiates the two regimes where oscillations around the fixed point are underdamped ($\zeta < 1$) or overdamped ($\zeta > 1$). The transition between these regimes ($\zeta = 1$) occurs when

$$f'(V^*) = \frac{RC}{L} \pm 2\sqrt{\frac{C}{L}}. \quad (4.9)$$

This implies that oscillations around the fixed point only occur when that fixed point falls in the valley of the IV curve, where $f'(V)$ is low, and that outside that small range, the state will exponentially approach the fixed point with no 'overshoot'.

Away from the fixed point, the nonlinearities (i.e. in the IV curve) become significant and methods of linear analysis cannot be used. Additionally, the equations of motion (equations 4.1 and 4.2) do not have analytic solutions, so must be solved by numerical methods. Some approximations can be made however, and the different groups of trajectories that occur can be described qualitatively. Figure 4.1.5 show the wider orbits ("relaxation oscillations") of the RTD around the IV curve when biased at $0.65V$, for different values of capacitance and inductance (near the default values). Since the fixed point for this bias value is inside the NDR, the stable state of the system is to orbit the IV curve with frequency determined by the capacitance and inductance.

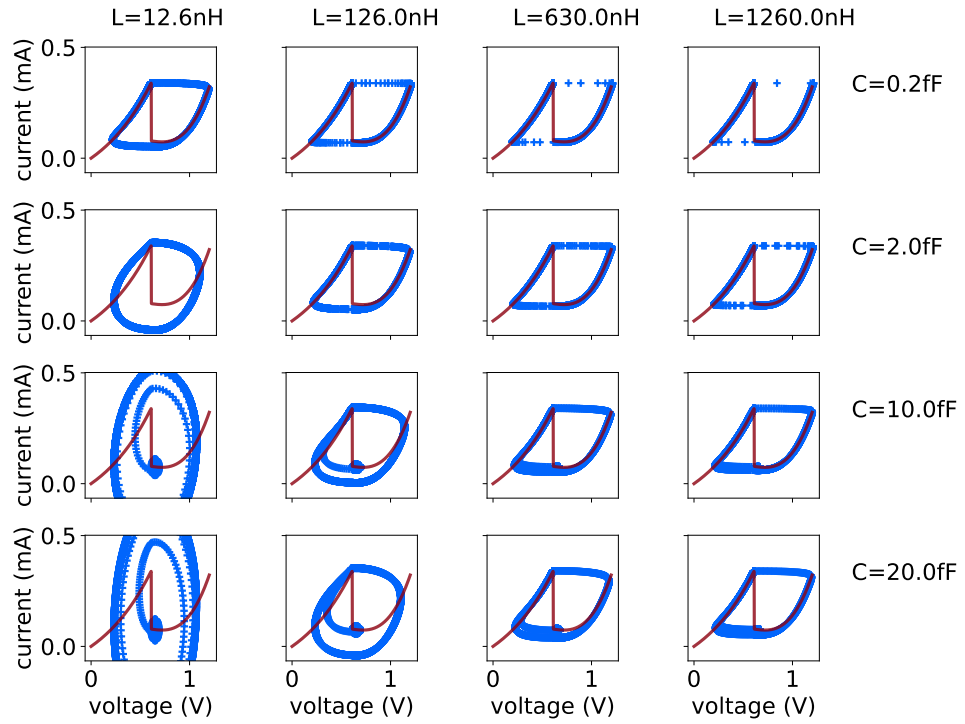


Figure 4.1.5: Orbits of the RTD circuit when biased at $0.65V$ for different values of capacitance and inductance.

The orbits can be described by four physical "phases", shown in figure 4.1.6. These phases can be explained by considering the movement of current carrying charges through the system: The voltage drop across the circuit's resistance and inductance, and the RTD voltage V must sum to the power supply voltage; and any current not flowing through the RTD (determined by $f(V)$) will accumulate (or discharge) charge on the circuit's capacitance.

- Phase 1: The voltage across the RTD is lower than the supply, so the current begins to increase. The RTD voltage also increases, as the current that can pass through the RTD is limited by $f(V)$ and so accumulates charge on the circuit's capacitance.
- Phase 2: The voltage across the RTD exceeds the voltage of resonant tunnelling, and the current the RTD can support rapidly drops. The excess current, which cannot change at the same rate due to the inductance, causes rapid accumulation

of charge and the voltage across the RTD quickly rises.

- Phase 3: The RTD voltage is now higher than the supply voltage, so the voltage across the resistive and inductive components is reversed, decreasing the current. Correspondingly, the capacitance discharges, decreasing the voltage across the RTD.
- Phase 4: Once the voltage reaches the resonant tunnelling point, the last accumulated charges are quickly discharged, dropping the RTD voltage.

Since these phases are a sequence of charging and discharging of the circuit's capacitive and inductive components, the values of R , L and C strongly determine how quickly the voltage and current can change, and so the speed of each phase of the oscillation. These parameters have the following influence:

- Resistance R : This determines where the I nullcline crosses the V nullcline, and so the position of the fixed point.
- Capacitance C : The capacitance occurs due to parasitics in the cables and connecting components, and due to the build up of charge on the barriers of the RTD[121]. This determines the voltage across the RTD that occurs due to charge build-up, and so the speed of phases (2) and (4).
- Inductance L : Inductance also occurs due to parasitics in cables and components such as power splitters and bias-tees. Since this affects how quickly current in the circuit can change the speed of phases (1) and (3) are dominated mostly by inductive effects.

Table 4.2 gives the value of R , L and C used by default for simulating the RTD circuit.

Comparison of this model and experimental[120, 122] results show that, given well fitted values for the $f(V)$ curve and values of R , L and C , the model is a good fit to reality, and so many experiments can be validated or designed using numerical simulation of this model

Parameter	Value
R	0.1Ω
L	126 nH
C	2 pF

Table 4.2: Values of resistance (R), capacitance (C) and inductance (L) used for simulating ultrafast RTD based neurons.

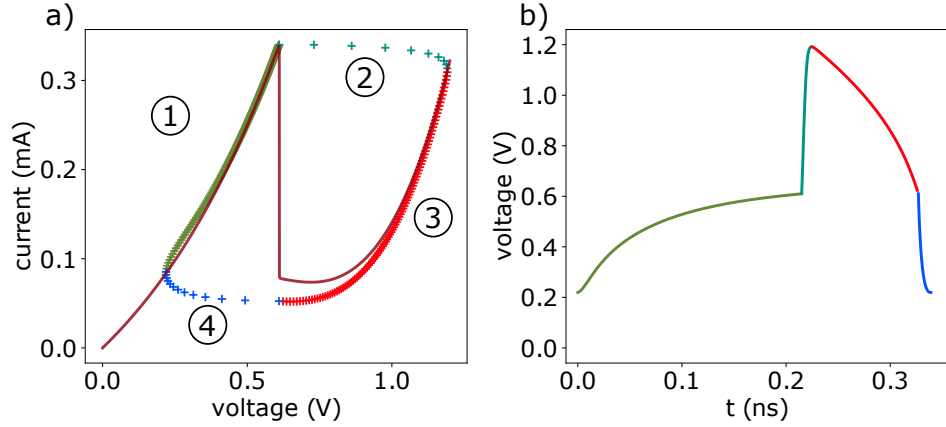


Figure 4.1.6: Phases of a relaxation oscillation of the RTD circuit biased at 0.65V . a) 1: Low voltage across the RTD causes the current in the circuit to increase, with the rate limited by the circuit inductance. Current that cannot pass through the RTD accumulates on the circuit capacitance, increasing the voltage across the barrier. 2) The voltage across the RTD passes the resonant tunnelling point, quickly reducing the amount of current that can pass through the barriers. The current that remains flowing (as a result of the inductance) accumulates on the capacitance of the circuit, rapidly increasing the RTD voltage. 3) The voltage across the RTD is now higher than the power supply voltage, and stage 1 is reversed. The capacitance discharges slowly through the non-resonant tunnelling channels of the RTD, reducing the current and RTD voltage. 4) The RTD voltage reaches the resonant tunnelling point, quickly discharging any remaining capacitive charge. b) The Voltage against time across the RTD, coloured to show the same phases as in (a). The phases 1) and 3) represent charging and discharging the capacitance are slow, while the phases associated with the effects of resonant tunnelling (2) and 4)) are fast.

4.1.4 Optically Coupled Resonant Tunnelling Diodes

The RTD circuit explored up until here has been entirely electronic, with the RTD state being described by the voltage and current, and the only control parameter being the input bias V_0 . In order to use this with photonics and light-based signals, the system must be extended to incorporate conversions from light to electrical signal and vice-versa. Numerous methods are available for controlling electronics with light (e.g. photodetectors, and in fact RTDs can be manufactured to be light-sensitive themselves), and for converting electronic signals to light - typically modulating the amplitude of light, so that it can go on to control further RTDs, or other photonic devices.

RTDs can be fabricated with transparent layers above the barrier structure that allow the RTD to respond directly to incident light[112]. In these devices, the layers above the barrier structure are made of a material transparent to the wavelength of light that can be absorbed by the photosensitive layer, which is fabricated below the barriers[113]. Light incident on this layer then excites conducting electron-hole pairs. Under low bias voltage, most of these photo-excited carriers quickly recombine without contributing to the current in the circuit, as the electric field across the absorbing layer is not strong enough to cause large movements of the carriers. At higher voltages, the electrons are carried away, contributing to the circuit current, and the holes accumulate at the barrier. This accumulation increases the voltage drop across the barriers, and so the electron tunnelling probability, shifting the IV curve to slightly lower voltages.

These effects can be modelled simply by adding the photocurrent term to the IV function $f(V)$ in the dynamical equations:

$$C\dot{V} = I - f(V) - \kappa S_0(t) \quad (4.10)$$

where κ represents the conversion efficiency from optical power to current, and S_0 is the optical power falling on the RTD.

An alternative to the directly photosensitive RTD is to drive the bias of the RTD circuit using an externally connected device, such as a photodiode. This allows the use of sensors with higher bandwidth than the RTD, sensitivity to different frequency

bands, or with specific functions such as balanced photodetectors.

In order to couple the RTD to the optical domain, the RTD circuit can simply be used as the driving component of any optically emitting system, such as an LED or laser diode. A two-level laser diode can be modelled [123] using rate equations describing carrier inversion number N and photon number S :

$$\dot{N} = N_{in} - (\gamma_l + \gamma_m + \gamma_{nr}) N - \gamma_m (N - N_0) S + \sqrt{\gamma_m N S} \xi(t) \quad (4.11)$$

$$\dot{S} = \left(\gamma_m (N - N_0) - \frac{1}{\tau_{ph}} \right) S + \gamma_m N \quad (4.12)$$

This models a laser with linear gain $G = \gamma_m(N - N_0)$, and transparency number N_0 . The rate coefficients γ_m , γ_l and γ_{nr} represent the rates of energy emission via spontaneous emission into the lasing mode, leaky radiative decay (into non-lasing modes) and non-radiative losses, respectively. τ_{ph} represents the mean free lifetime of a photon in the cavity, in the range of picoseconds to femtoseconds. All sources of noise are collected into the white noise function $\xi(t)$.

The laser model is driven by the N_{in} term, representing the pumping of carriers to increase the inversion count, which can be driven using either the voltage of the RTD, or the current. The current controlled term is written $N_{in} = \frac{J + \eta I}{q_e}$, while the voltage controlled term is $N_{in} = \frac{J + \eta V / R_0}{q_e}$, with J being an additive bias current and R_0 representing the impedance of the laser driver, converting the voltage into a current modulation.

Table 4.3 gives the values of the photodetecting RTD sensitivity, and laser diode parameters used for modelling of RTDs in this work.

This laser model shows a very typical current-intensity (L-I) curve: below a threshold current J_{thr} , the power output is insignificant, while above threshold, the power increases linearly with current. For the laser modelled with these values, the threshold current is

$$J_{thr} = q_e (\gamma_l + \gamma_m + \gamma_{nr}) \left(N_0 + \frac{1}{\gamma_m \tau_{ph}} \right), \quad (4.13)$$

Parameter	Value
κ	1×10^{-7}
N_0	5×10^5
γ_l	10^9
γ_m	10^7
γ_{nr}	2×10^9
τ_{ph}	0.5 ps
η	1

Table 4.3: Parameters used to model a simple two-level diode laser.

which evaluates to $337 \mu A$, and the output power at equilibrium, when the current J is above the threshold is

$$S(J) = \tau_{ph} \left(\frac{J}{q_e} - (\gamma_l + \gamma_m + \gamma_{nr}) \left(N_0 + \frac{1}{\gamma_m \tau_{ph}} \right) \right) \quad (4.14)$$

The bias current term in equation 4.11 allows the laser to respond quickly to changes in the driving value from the RTD, by biasing the laser just below its threshold, any increase in the driving signal is immediately converted into an increase in optical output.

Figure 4.1.7 shows each component of the RTD model during a spike triggered by a 100ps long input pulse of optical power.

4.1.5 Numerical Simulation of RTDs

The differential equations describing the RTD circuit (and coupled laser diode) are highly nonlinear, and do not have any exact solutions (except for the case of trivial inputs such as $V_0(t) = 0$ and $S_0(t) = 0$). Numerical solutions that are very good approximations, however are possible to find by computer integration.

Numerical integration of the purely electric RTD circuit given by equations 4.1 and 4.2 can be done accurately and quickly using established methods, such as the fourth-fifth order Runge-Kutta method, which is easily accessed via SciPy's[124] 'solve_ivp' function. The full photodetecting-RTD-laser diode model however does not work well with the explicit integration method due to the $1/\tau_{ph}$ term that leads to the laser equations being 'stiff' and requiring an intractably small time step to be integrated. Instead, an implicit integration method can be used.

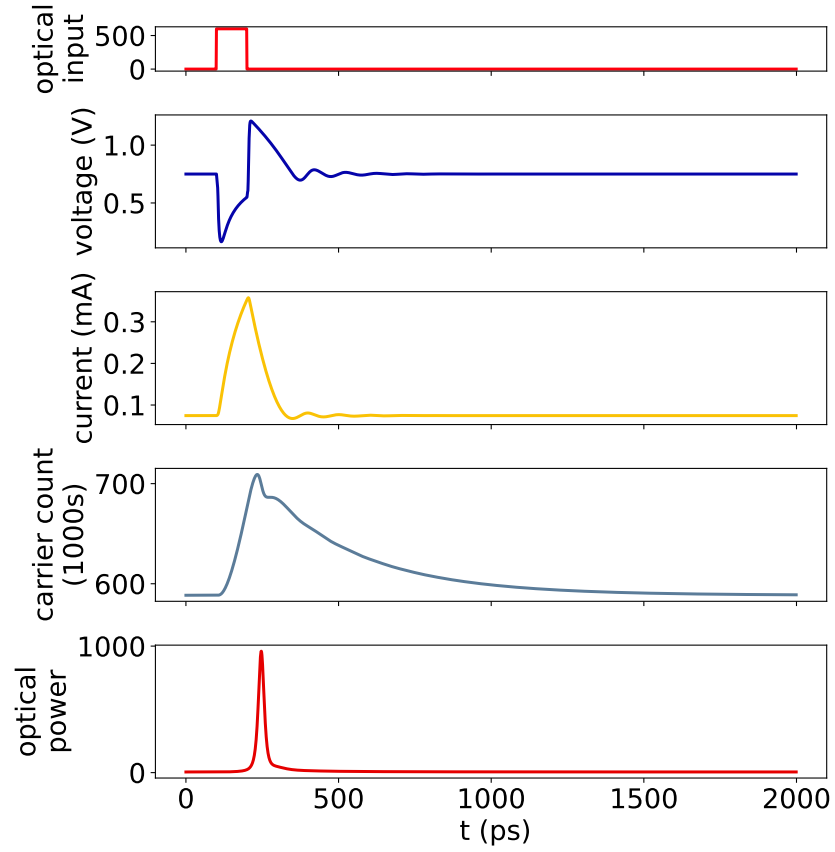


Figure 4.1.7: Voltage, Current, Carrier inversion count, and laser output during a single spike, triggered by a 100ps long optical input. The RTD current drives an increase in the carrier inversion, which in turn drives an increase in the output of the laser.

An implicit Gauss-Legendre integration method[125] of the differential equation $\dot{y} = f(t, y)$ takes the form:

$$t[n + 1] = t[n] + \Delta t \quad (4.15)$$

$$y[n + 1] = y[n] + \Delta t \sum_{i=1}^s b_i k_i \quad (4.16)$$

$$k_i = f \left(t[n] + c_i \Delta t, y[n] + \Delta t \sum_{j=1}^i a_{ij} k_j \right) \quad (4.17)$$

Solving the implicit equations for k_i is done by Newton's method.

For integration of the RTD-laser circuit, the number of stages used was $s = 2$, giving a fourth order integrator, allowing accurate solutions with a time step as high as $1ps$. The values for the coefficients a_{ij} , b_i and c_i are given by[125]:

$$a_{ij} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} - \frac{\sqrt{3}}{6} \\ \frac{1}{4} + \frac{\sqrt{3}}{6} & \frac{1}{4} \end{bmatrix} \quad (4.18)$$

$$b_i = \begin{bmatrix} \frac{1}{2}, \frac{1}{2} \end{bmatrix} \quad (4.19)$$

$$c_i = \begin{bmatrix} \frac{1}{2} - \frac{\sqrt{3}}{6}, \frac{1}{2} + \frac{\sqrt{3}}{6} \end{bmatrix} \quad (4.20)$$

A custom solver implementing this method was created that allowed the use of controllable, possibly recurrent, connections between simulated RTD neurons, that would be difficult to implement otherwise using existing software.

4.2 Spiking Neurons Based on Resonant Tunnelling Diodes

The RTD under constant bias has two stable states: the stable fixed point, and the limit cycle around the IV curve. Transition from the fixed point to the limit cycle can be achieved by perturbing the RTD (with varying bias voltage or optical input) such that the system moves onto the limit cycle, returning to the fixed point. This ability to trigger an excursion onto the limit cycle is known as excitability.

These spikes can be triggered by biasing the RTD at a stable point near to the NDR, and perturbing such that the flow from what was the fixed point brings the circuit onto the limit cycle. When the RTD is biased in the valley, this means reducing the voltage, and in the peak means increasing the voltage. Spikes occurring from the peak and valley appear inverted, since the limit cycle always follows the same direction around the IV curve.

This section will demonstrate the excitable spiking and the different ways the RTD neuron can respond to input signals, electrically and optically.

4.2.1 Thresholding

Since the fixed point of the RTD circuit has a region in which all trajectories immediately return (typically referred to as the 'basin of attraction'), for a perturbation to cause the RTD to spike must be sufficiently strong, so that the circuit has time to escape the fixed point and enter the limit cycle.

Figure 4.2.1 shows the trajectory flows when biased in the IV-curve valley ($0.75V$) and peak ($0.6V$), and the time taken for the circuit to reach the relevant fixed point when starting at each V, I value. The trajectory time highlights the spiking threshold: during a finite duration perturbation, if the circuit state crosses the threshold line (plotted in red dashes), the only path it can take back to the original fixed point is to travel around the IV curve on the spiking trajectory, otherwise, it returns directly.

The implication of this is that for a pulsed perturbation to bias voltage of fixed duration, the amplitude must be sufficiently high so that the circuit can move to the boundary within the time of the perturbation. Figure 4.2.2 shows the response of the RTD (biased at $0.75V$) to pulsed decreases in the bias voltage, for different amplitude and duration of the pulses. When the bias is decreased, the flow at the location of the fixed point is changed to move towards the new fixed point, bringing the circuit towards the spike threshold, with further decreases in bias leading to faster flow. If the bias is returned to its original value before the circuit crosses the threshold, it will simply return to the fixed point with no spike (though with a few small oscillations), otherwise a spike will occur. Equivalently, a low-amplitude pulse can cause a spike if the pulse is long enough, such that the circuit can cross the threshold even if the flow is not as fast.

The threshold can be effectively controlled by adjusting the bias voltage of the RTD, since while the fixed point is dependent entirely on the bias, the boundary between the spiking and non-spiking trajectories does not change significantly. When the bias is set such that the fixed point is far from the spiking boundary, the amplitude of pulse required to reach the boundary becomes proportionally higher.

Thresholding applies similarly to the use of optical pulses for triggering spikes, as shown in figure 4.2.3, where 20 ps optical pulses only trigger spikes when the amplitude

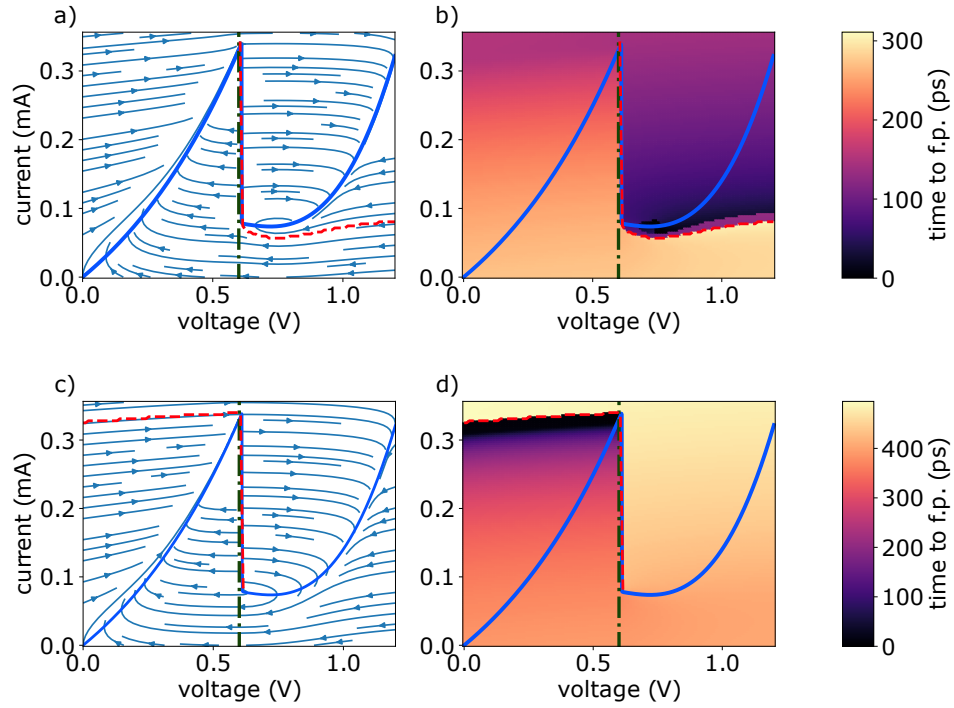


Figure 4.2.1: Trajectories of the RTD circuit towards the fixed point when biased in the valley ($0.75V$, a-b) and in the peak ($0.6V$, c-d). a,c) Trajectory flow towards the fixed point. b,d) Time taken (in picoseconds) to reach the fixed point from each initial V, I values. Spiking trajectories are those that travel back to the fixed point around the IV curve, and so take a long ($> 300ps$) time. The boundary between the long time and short time trajectories, plotted as a red dashed curve, represents the spike threshold: if the circuit moves beyond this curve, it is forced onto the spike trajectory.

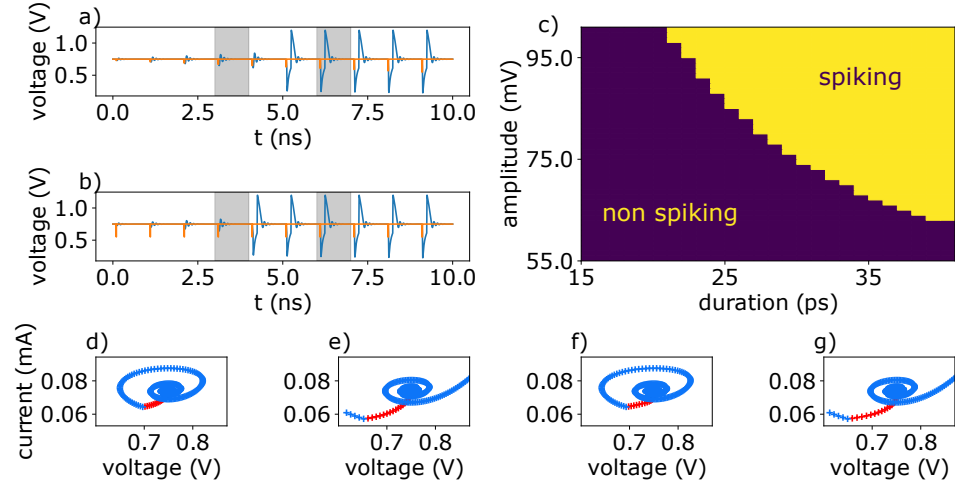


Figure 4.2.2: Threshold mechanism for spiking given square pulse modulations of the bias voltage, when biased at $V_0 = 0.75V$. a) 20 ps pulses of increasing amplitude are used to perturb the RTD. Spikes are only emitted when the amplitude is high enough (in this case when the pulse is at least 0.12V. b) rectangular pulses of constant amplitude (0.2V), with increasing duration. only when the duration of the pulse is longer than 10 ps does the RTD spike. c) map of which combinations of pulse amplitude and duration cause a spike. Higher amplitude and longer duration both lead to spiking. d-e) The trajectories followed by the RTD for a sub-threshold and super-threshold amplitude pulse (shaded regions in a)). The red trajectory indicates the time when the perturbation was on. f-g) The trajectories followed by the RTD for a sub-threshold and super-threshold duration pulse (shaded regions in b)). The red trajectory indicates the time when the perturbation was on.

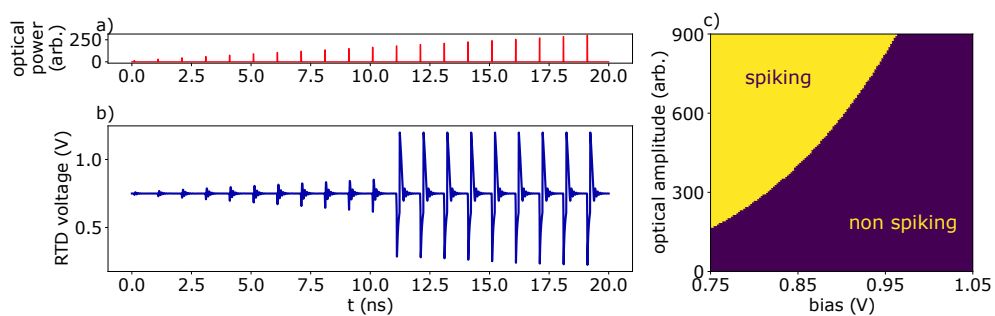


Figure 4.2.3: Thresholding of optical pulses. The RTD is biased in the valley, at 0.75V, and triggered with 20 ps optical pulses. a) Optical pulses 20 ps long with increasing amplitude. b) Voltage response of the RTD given the input in a). c) Parameter map showing the dependence of the amplitude on bias voltage. As the bias is increased, away from the IV valley, the amplitude required to trigger a spike also increases.

is high enough, with the threshold tunable by bias voltage.

4.2.2 Refractoriness

After a spike has been triggered, the circuit must return to the fixed point (or nearby) before another spike can be triggered again. Clearly, when the circuit is undergoing the orbit around the IV curve in the duration of a spike, a spike cannot be 'triggered', however there is a short period of time (in the final phase of the cycle) in which a spike can be triggered, albeit with a higher threshold. This inhibitory feature is known as the "refractory period", and occurs commonly in biological neurons (e.g. the neuron must replenish its neurotransmitters before spiking again) and other excitable systems.

This is demonstrated in Figure 4.2.4, where the RTD is perturbed by super-threshold pulses (individually able to trigger a spike from the resting state) that occur close together, within several hundred picoseconds.

4.2.3 Integrate and Fire

While the boundary between the spiking and no spiking trajectories gives a hard threshold for the size of a single perturbation, any input that causes the RTD to follow a trajectory that reaches the boundary can cause a spike. As explained previously, the dynamics of the RTD around its fixed point is oscillatory, with a damping param-

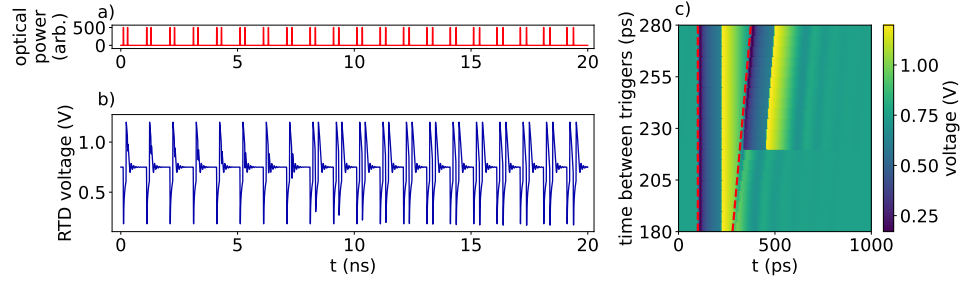


Figure 4.2.4: Demonstration of the refractory effect of spikes triggered by 20 ps pulse of optical power, all of equal amplitude, above the threshold amplitude. Pulses arrive with time between the start of the pulses from 180 ps up to 280 ps . The first pulse always triggers a spike, as it is well over threshold, however the second pulse can only trigger a spike after the refractory period has passed. a) The optical input to the RTD. b) the response of the RTD. c) time map of the spikes emitted in b). The red dashed lines mark when the pulses arrive.

ter dependent on the local gradient of the IV curve, with high gradients leading to overdamped relaxations, exponentially approaching the fixed point.

This slow decaying response to input perturbations allows the RTD to "integrate" input perturbations over a short time, and fire a spike in response to the the integrated perturbation. This mode of response is used in a model of biological neuron responses known as the "leaky-integrate-and-fire" model, based on the idea of ion channels in the neuron's membrane that leak charge.

Figure 4.2.5 shows the response of an RTD neuron biased near the IV peak (at 0.55 V) to a sequence of rectangular pulsed increases in bias. Each individual pulse on its own is 20 ps long and 0.3 V in amplitude, not enough to trigger a spike. The first pair of pulses are separated by 500 ps , and the RTD's responses do not overlap. The second pair arrive separated by 200 ps . In this case, the response of the RTD to the first pulse has not entirely decayed, and so the effect of the second pulse pushes the circuit closer to the threshold.

The sub-threshold dynamics of the RTD can be approximated by equation 4.21, where the input $\tilde{V}_0$ is the modulation of bias around the constant background value. This takes the form of convolution with the exponentially decaying impulse response, as in leaky-integrate-and-fire models.

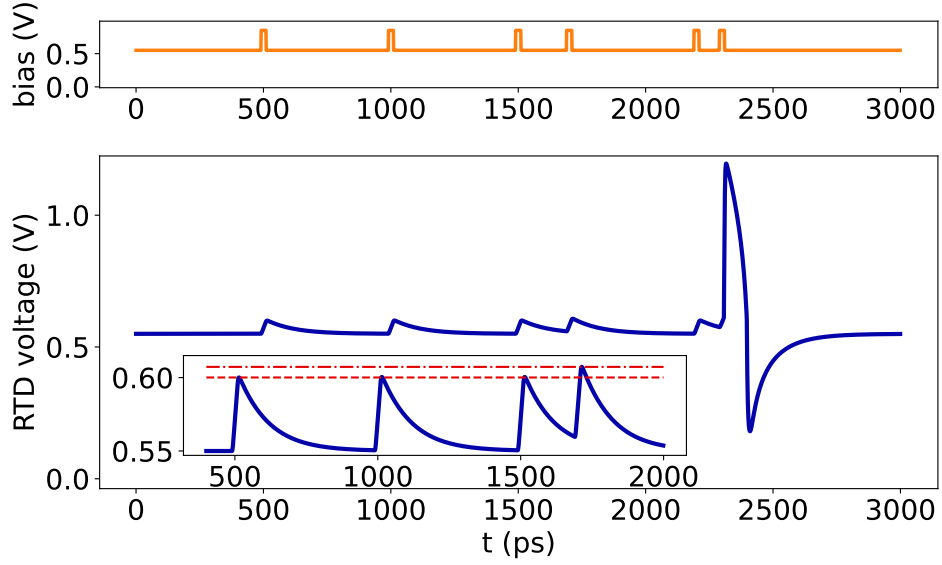


Figure 4.2.5: Integration of two 20 *ps* rectangular voltage pulses to cause a spike. The RTD is biased at 0.55V, near the peak of the IV curve. Positive perturbations to the bias voltage bring the circuit closer to the threshold, however if a spike is not started before the perturbation ends, the voltage will decay exponentially back to the resting point. A second pulse arriving before the response totally decays brings the voltage to a higher point, closer to the threshold (inset). If the two pulses are near enough, the combined response is enough to cross the threshold and the RTD spikes.

$$V(t) = \int_0^\infty \tilde{V}_0(t - \tau) e^{-\omega(\zeta - \sqrt{\zeta^2 - 1})\tau} d\tau \quad (4.21)$$

4.2.4 Resonate and Fire

Equivalently, when the IV curve gradient is low near the fixed point, the oscillations become underdamped, and after any perturbation the RTD circuit will return to the fixed point via several oscillations. This can be approximated by convolution with the exponentially decaying sinusoidal impulse response:

$$V(t) = \int_0^\infty \tilde{V}_0(t - \tau) e^{-\omega_0 \zeta \tau} \sin(\sqrt{1 - \zeta^2} \omega_0 \tau) d\tau \quad (4.22)$$

Since the frequency ω_0 and damping ζ depend on the gradient of the IV curve, the bias point of the RTD will determine the resonant frequency and impedance, with bias

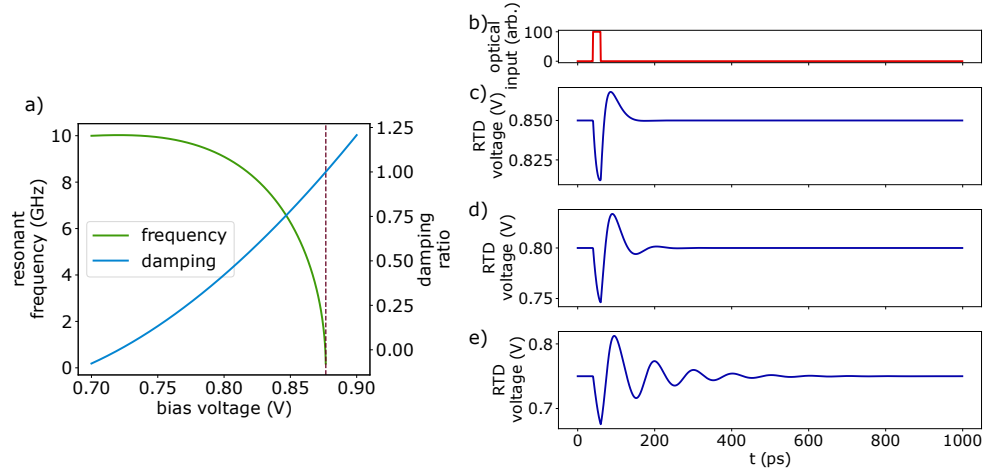


Figure 4.2.6: The resonance response of the RTD circuit to small perturbations when biased in the valley. a) dependence of damping ζ and resonant frequency $\omega_0\sqrt{1-\zeta^2}$ on bias. At a biases above 0.875V (marked by the dashed line), the system is overdamped, and so the response no longer oscillates. b) A 20ps pulse of optical power applied to and RTD, with the responses at different bias voltages shown in c (0.85V), d (0.8V) and e (0.75V)

increasing (away from the NDR) leading to higher damping and a reduction in resonant frequency, as demonstrated in figure 4.2.6.

As with the integrate-and-fire mechanism, this allows the RTD to fire spikes in response to a sequence of input perturbations that collectively cross the threshold, even if any individual perturbations would not.

Figure 4.2.7 shows the result of sequences of short voltage perturbations of different polarity (increase v.s. decrease) and separation time applied to an RTD biased in the valley - where $f'(V)$ is low enough that the circuit will oscillate. Individually, each of these voltage pulses do not trigger spikes, however, where their responses align constructively, a spike can be triggered.

4.3 Experimental investigation of RTD neurons

This section will describe the experimental setups used to investigate the photonic-electronic Resonant Tunnelling Diode based circuits. For the experiments, the circuit was assembled using RTD devices connected to commercially available off-the-shelf

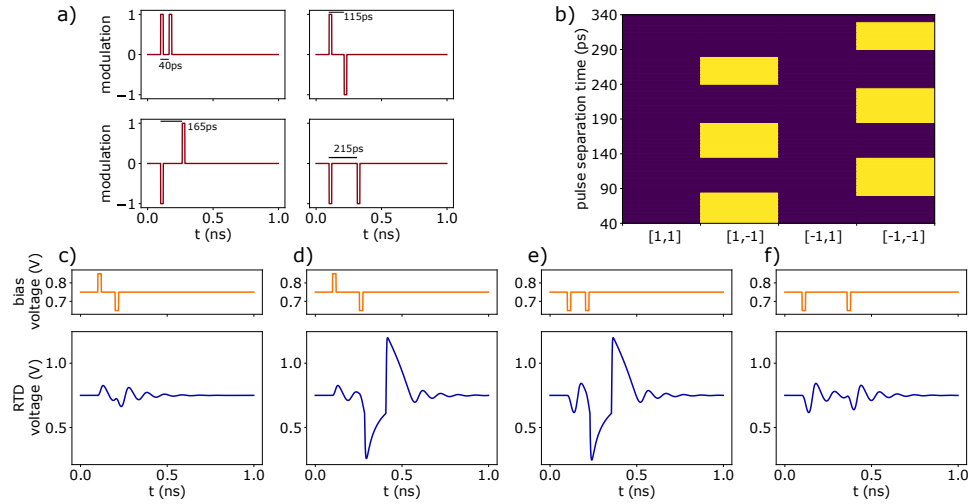


Figure 4.2.7: Pairs of voltage pulses can trigger spikes when their responses collectively cross the spike threshold. a) Examples of the modulations used. The pairs of pulses were either the same polarity or opposite, with varying separations between the pulses. b) Whether the RTD spiked (yellow) or not (purple) for combinations of positive and negative voltage pulses at certain separations. c-f) Examples of the RTD response to voltage pulses on and off resonance. pulse separations are c: 100ps, d: 155ps, e: 105ps, f: 260ps

electronic and optical equipment in a modular fashion such that parts could be disconnected and added easily.

The photonic-electronic RTDs were fabricated at the University of Glasgow[114], and were built from nanometer sized layers of semiconductor, as described in table 4.4. These devices incorporated a window in the top layers (a hole in the top electrical contact) that allows the device to respond directly to optical signals, as described in the modelling section previously.

Figure 4.3.1 shows a schematic describing the connections between experimental equipment used to control and probe Resonant Tunnelling Diode neurons. The RTD device was probed with a custom built station, using a GSG probe (GGB Picoprobe Model 40A) to connect to the RTD chip, and a Signatone S-M40 micropositioner stage. The GSG probe was connected to a power splitter that allowed input voltage signals to reach the RTD, and for the RTDs response to be transmitted forward to an oscilloscope (a Rohde&Schwartz RTP 16GHz real-time oscilloscope). The input voltage to the RTD

Layer	Material	Thickness (nm)
Contact	InGaAs	1
Collector	InAlAs	1
Spacer	InAlGaAs	250
	InGaAs	2
Barrier	AlAs	1.7
Well	InGaAs	5.7
Barrier	AlAs	1.7
	InGaAs	2
Spacer	InGaAs	20
Emitter	InAlAs	1
Contact	InGaAs	5

Table 4.4: Material composition of the photonic-electronic RTD device. This was designed and fabricated by our project collaborators at the University of Glasgow (Dr Qusay Al-Taai, Prof Edward Wasige). In bold are the layers that form the double-barrier quantum well, responsible for the nonlinear dynamics of the RTD.

was controlled by a power supply that set the bias voltage, and an arbitrary wave generator (12GSa/s Keysight M8190A). The outputs of the power supply and wave generator were combined using a 12 *GHz* bias-tee (Inmet 8800SMF1-12), which was then connected via the DC+RF port to the power splitter, and thus to the RTD. Each of these components (including connecting SMA cables and their couplers) contribute a parasitic inductance, capacitance and resistance to the circuit, which is modelled above by the collective terms R , L and C . Thus, even in this circuit, where there is no explicit capacitor or inductor present, oscillations and spikes can occur.

The AWG was also used to control the optical input to the RTD, by controlling a Mach-Zehnder modulator (MZM) that modulates the intensity of light from a tunable laser. The tunable lasers used were a Santec TSL-210 for operation at 1310 *nm*, and a Santec Wsl-110 operating at 1546 *nm*. Injection in to the RTD was done using a lens-ended fibre. Figure 4.3.2 shows the experimentally obtained IV curve of one such device, under dark conditions, and under illumination with 1 *mW* of 1310 *nm* light. This is collected by setting the bias voltage, and measuring the time averaged current in the circuit for each value of bias, resulting in the recording current at the stable point of the RTDs operation (or mean current during a spike when biased in the NDR). The illuminated IV curve shows the increase in current (as a direct result of the

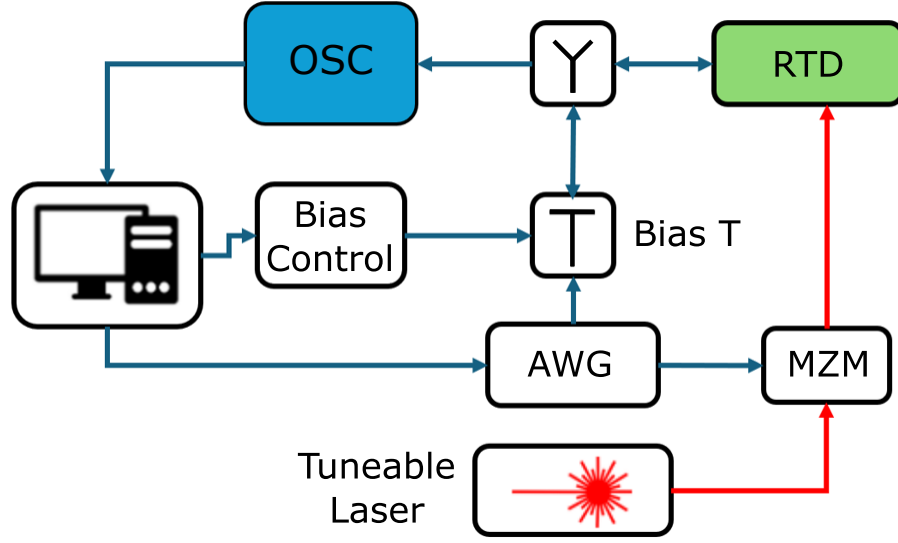


Figure 4.3.1: Schematic of the experimental setup used to probe the photonic-electronic RTD neuron. The bias voltage to the RTD is set by a power supply (bias control), and modulated with signals from an arbitrary wave generator (AWG), combined by the bias-tee. The optical power landing on the RTD was modulated via an MZM. The electrical output of the RTD is measured and recorded using a real-time oscilloscope (OSC)

photocurrent) and a slight shift of the peak and valley to lower voltages).

Very recent works from our group have demonstrated neuron-like spiking in these devices, triggered by input optical and electrical signals. The features demonstrated previously in simulation, such as thresholding, integration, and refractoriness have also been shown experimentally[111, 113, 114, 126].

Figure 4.3.3 demonstrates the triggering of a spike in an RTD using a short $5ns$ pulse of optical power. Panel a) shows the recording of the intensity modulation sent to the RTD, with the pulse occurring at $20ns$. The mean power of this input was adjusted such that $1.8mW$ was falling on the RTD's photosensitive layer, which was enough to trigger a spike. Panel b) shows the resulting $\sim 150ns$ long spike in the voltage across the RTD, as measured by the oscilloscope. The RTD was biased at $0.74V$, inside the valley near to the edge of the NDR, allowing the increase in light to trigger the spike.

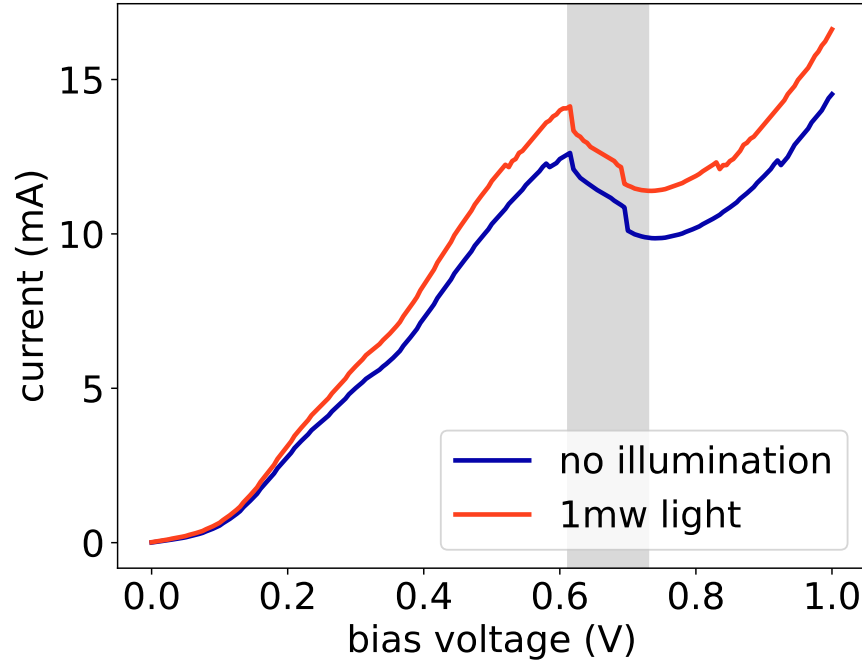


Figure 4.3.2: Experimentally measured IV curve of a photonic-electronic RTD, under darkness and illumination of

1 mW

of 1300nm light. Shaded in grey is the region of negative differential resistance.

4.4 Summary

This chapter introduced the resonant tunnelling diode, a semiconductor device formed by a double-barrier quantum well. These diodes feature a nonlinear, N-shaped conductivity curve, resulting from the varying efficiency of quantum tunnelling across the well with voltage. RTDs were shown as spiking neurons, exhibiting integrate-and-fire and resonate-and-fire behaviour, in response to electrical and optical modulation.

In the following chapter, this setup will be used to realise experimentally a neuromorphic use-case of a photonic-electronic RTD neuron that is also demonstrated in simulation, showing good agreement between experimental and theoretical results.

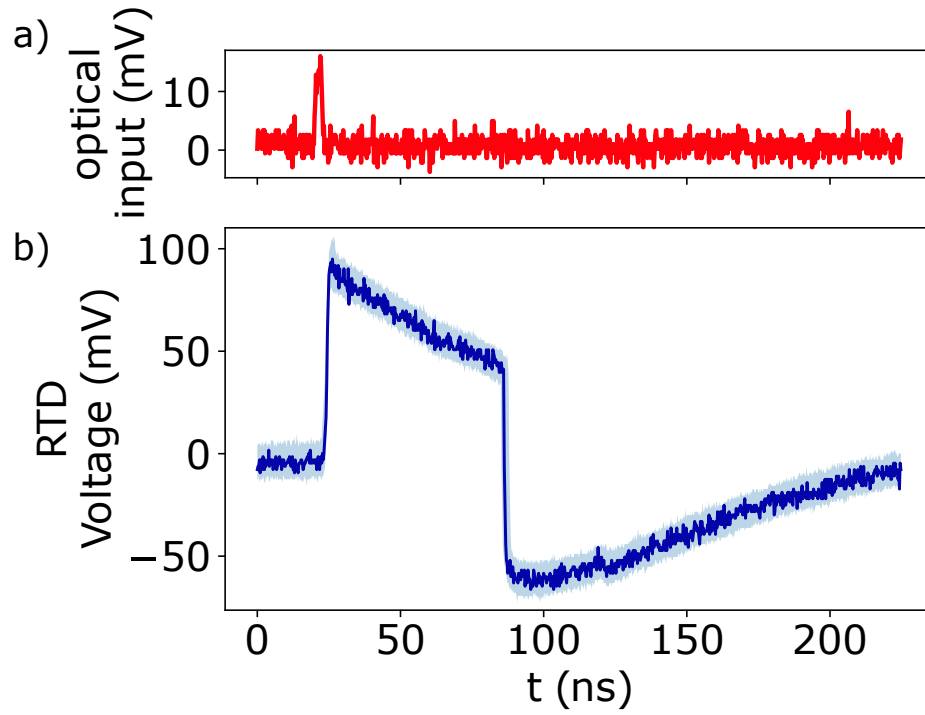


Figure 4.3.3: Recording of the voltage across an RTD during a spike triggered by a short ($5ns$) optical pulse. The RTD was biased in the valley, at $0.74V$, allowing spikes to be triggered by increasing pulses of optical power. a) shows the recording of the optical modulation. The mean power of light falling on the RTD was set to $1.8mW$ for reliable spike triggering. b) shows the RTD voltage. The shading shows the maximum and minimum voltage over 100 repetitions, demonstrating the consistency of the spiking response.

Chapter 5

Neuromorphic Processing Using RTDs

This chapter will present the use of these photonic-electronic RTD based spiking neurons for neuromorphic information processing and computing. Different architectures of RTD neuron based networks will be demonstrated for tasks such as processing time-based data, neuromorphic encoding and storage, and typical AI tasks like classification.

This chapter reports simulation results produced with the theoretical model included in Chapter 4. This chapter also includes experimental results describing the use of photonic-electronic RTD spiking neurons applied to a time-series processing task, namely the fast detection of rising edges in time-series data, with good agreement between experimental and numerical findings.

Relevant Publications

- D. Owen-Newns, J. Robertson, G. Donati, Q. Al-Taai, J. Figueiredo, E. Wasige, K. Lüdge, B. Romeira, A. Hurtado, "Neuromorphic Photonic Processing and Memory with Spiking Resonant Tunnelling Diode Neurons and Neural Networks," *Advanced Intelligent Systems*, vol. 0, p. 202500800, 2026
- G. Donati, D. Owen-Newns, J. Robertson, X. Porte, E. Malysheva, J. Figueiredo, B. Romeira, V. Dolores-Calzadilla, A. Hurtado, "*Spiking Flip-Flop Memory in*

Resonant Tunneling Diode Neurons,” Phys. Rev. Lett., vol. 133, p. 267301, 26 Dec. 2024.

5.1 Neuromorphic Processing with Single Photonic-Electronic RTD Neurons

5.1.1 Edge-Feature Detection in Time-Series

As described in Chapter 4, an optically sensitive RTD biased in the IV curve valley region can be made to fire a spike by an incoming perturbation (i.e. an optical/electrical pulse) with strength above the device’s excitation threshold. This is shown graphically in Figure 5.1.1 for fixed-duration input optical pulses. The parameter map shown in Figure 5.1.1(a) reveals combinations of applied bias voltages and input optical pulse amplitudes for which the RTD neuron will (or not) fire an excitable spike upon the arrival of a 20ps-long optical input, and figure 5.1.1(b) plots the numerically simulated output time-series of the RTD neuron when operated with bias voltages of 0.75 V and 0.85 V, for different optical input amplitudes (indicated on the map using blue and red symbols).

This optical-electrical dual modulation feature of the RTD neuron is used to perform event-based detection of rising edges in time-series inputs, by programming the RTD neuron to fire excitable spikes when the amplitude of an input signal increases by a sufficient amount over a short period of time, i.e. when the difference D_τ (according to Equation 5.1) exceeds some value.

$$D_\tau(t) = y(t) - y(t - \tau), \quad (5.1)$$

Raw time-series values $y(t)$ are combined with a delayed signal copy $y(t - \tau)$ to create $D_\tau(t)$, that when high can be setup to exceed the spike-firing threshold of an RTD neuron. To achieve this, the raw time-series value ($y(t)$) was encoded onto the amplitude of rectangular optical pulses, while the delayed time-series value ($y(t - \tau)$) was then used to control the bias voltage of the RTD, modulating dynamically the spiking

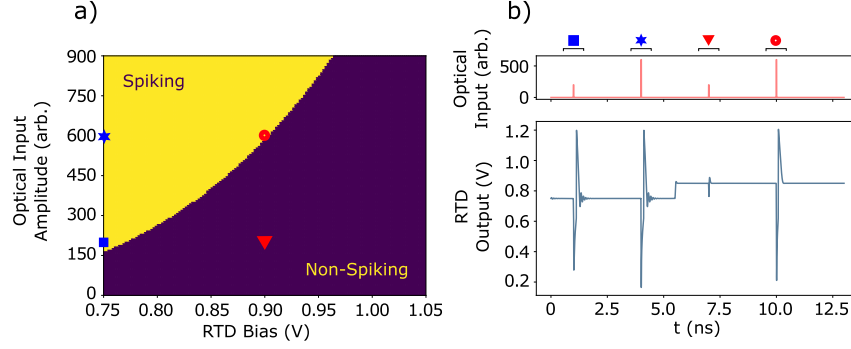


Figure 5.1.1: Varying the spike firing amplitude threshold of an input optical pulse by varying the bias voltage (a) Optical spike firing threshold amplitude against bias voltage. As the bias on the RTD is increased, the amplitude of an incoming optical pulse must be higher in order to trigger a spike response in the system (yellow pixels represent a spike being triggered, while purple represents no spike). (b) Simulated optical inputs of 20 ps-long pulses of different intensities injected into an RTD (top, red time traces). Simulated output voltage of an RTD neuron's response (bottom, blue time-series) when the device is biased at two voltages, 0.75 V and 0.85 V. Respective symbols indicate the location of input conditions on the map.

threshold of the RTD neuron. Figure 5.1.2(a) illustrates graphically the operating principle of the RTD event-based time-series edge detection system.

Figure 5.1.2(b) and 5.1.2(c) plot the simulated input time-series signal and the copy, delayed by 10 time steps, while figure 5.1.2(d) plots their respective difference. The optical input signal is created using 20 ps-long pulses repeated every 2 ns. The maximum amplitude of the optical input was 600 arbitrary units, as used in Figure 5.1.1(b), scaled by the input in figure 5.1.2(b). The delayed copy is created using a 0.75 V bias voltage and a modulation voltage of up to 0.15 V, creating a maximum bias voltage of 0.9 V during the peak of the delayed time-series signal (Figure 5.1.2(c)). This signal contains several regions that show the working principles of the edge detection system:

- 20 ns until 30 ns: The time-series increased and so optical pulses are now falling on the RTD, but until the 10 ns delay has passed, the bias remains low, allowing the optical pulses to trigger spikes
- 30 ns until 40 ns: The time-series value is still high, but now the delayed copy

is also high. The increased bias voltage, and so increased threshold means the optical pulses do not trigger spikes (but cause small oscillations). This represents the time-series having a high value, but without significant change over the delay time, and so no rising edge.

- 40 ns until 50 ns: While the bias remains high, the time-series has now decreased and so no spikes are triggered.
- 60 ns until 90 ns: The same process, however the absolute change in the time-series is smaller. The mapping from time-series value to optical pulse amplitude determines the size of the rising edge to be detected.
- 100 ns until 120 ns: The time-series value increases, but the change is too small to trigger spikes, even at low bias voltage.
- 120 ns onward: The further increase of the time-series means the optical pulses now again trigger spikes, even with the increased bias voltage from the step occurring at 100 ns

To further test the RTD neuron at this task, it was applied to the detection of rising edges in the a complex chaotic time-series, namely the Mackey-Glass (MG) time-series [106]. The latter follows a quasi-periodic pattern and has many regions with differing gradients, hence defining a much more complex rising edge detection task. Figure 5.1.3 shows numerical results of a multi-modal RTD spiking neuron tasked with the Mackey-Glass time-series rising edge detection task. Figure 5.1.3(a) shows the Mackey-Glass time-series signal analysed in the study. Figure 5.1.3(b) visualises the delay difference between the raw MG signal, and a copy delayed by 10 time steps. The regions shaded in grey indicate the rising edges that are targeted by the RTD spiking neuron.

For this case of analysis, the values of the generated MG time-series (featuring an MG tau of 17) were first normalised to the range 0 to 1. The raw MG optical input was then created with pulses up to an amplitude of 600 arbitrary units, widths of 20 ps, and separations of 500 ps. The delayed copy of the MG time-series was also electrically injected into the RTD neuron. A constant voltage of 0.75 V was modulated with up to

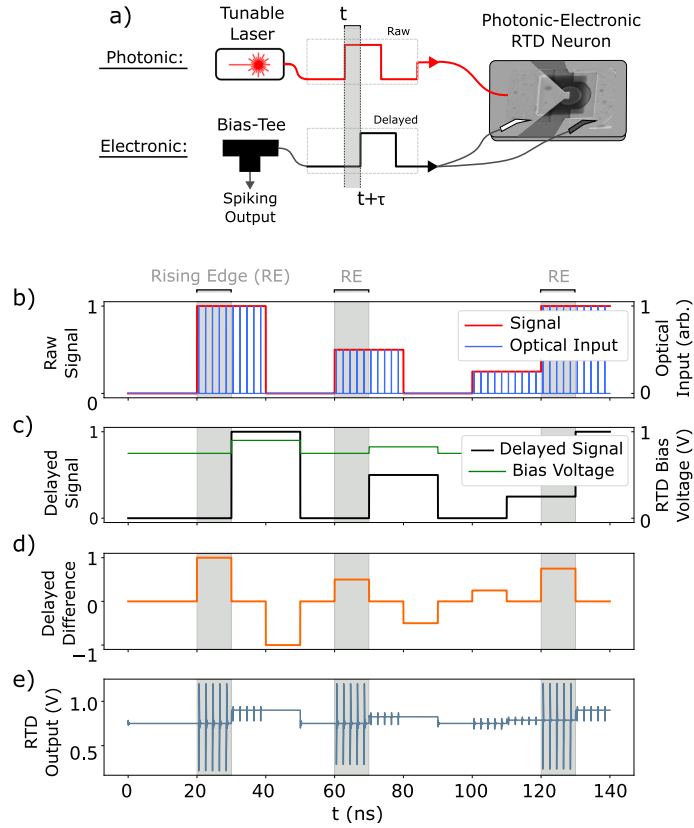


Figure 5.1.2: Operation principle of RTD event-based time-series rising edge detection. (a) Schematic of a photonic-electronic RTD neuron subject to simultaneous modulation of its optical input and bias voltage. (b) An example raw signal is simulated for the purpose of the demonstration (red line). The raw signal amplitude is encoded into short (20 ps) rectangular optical pulses (at 0.5 GHz) to optically modulate the RTD neuron (pink line). (c) A 10 ns-delayed copy of the raw time-series signal (black line) is also electrically encoded into the RTD by modulating its voltage bias. An amplitude modulation of up to 0.15 V is performed on-top of the 0.75 V constant bias to push the RTD operation point away from the valley region of its I-V curve (green line). (d) The calculated amplitude difference between the raw signal and its 10 ns-delayed copy (orange line). The delay difference visualises the simultaneous, counteracting, photonic-electronic inputs and reveals a clear threshold above which the RTD neuron will fire. (e) Calculated temporal output from an RTD subject to photonic-electronic modulation (blue line), demonstrating successful edge detection via fast spike-firing events (marked with grey shadow regions).

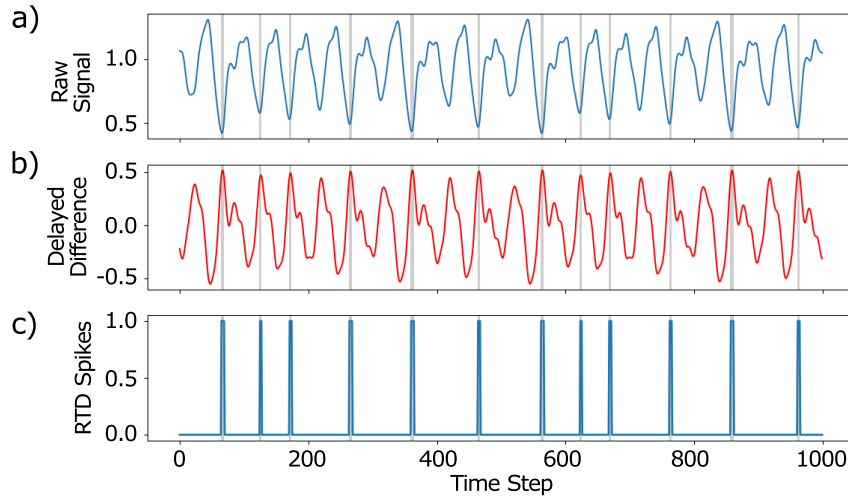


Figure 5.1.3: Rising edge detection of the Mackey-Glass time-series performed by a simulated RTD. a) The Mackey-Glass chaotic time-series signal. b) The calculated delayed difference between the raw MG signal and a copy delayed by 10 time steps. c) Time steps for which the RTD emitted a spike ('0' represents no spike, '1' represents a spike), showing that the dual modulation method can reliably detect rising edges. Highlighted in grey (a-c) are regions where edge detections are expected after exceeding a delayed difference of 0.45.

0.2 V for a maximum voltage of 1.05 V. Figure 5.1.3(c) plots the time steps in which the RTD neuron produced spike events ('1' spiking, '0' no spiking). These are overlaid by grey shaded regions, showing where edge detections are expected. A threshold of 0.45 delayed difference was found to detect the sharpest rising edges in the MG time-series (see Figure 5.1.3(b)). This threshold value (0.45) was then implemented empirically by correctly choosing the optical pulse amplitude and voltage modulation levels that together produced the matching spiking output. The speed of edges (i.e. shallow or steep) can be adjusted by changing the delay between the samples used for optical and electrical modulation. This threshold parameter is intrinsic to the signal used and we therefore do not rely on it to program the spiking threshold of the system. Once the RTD spiking threshold was correctly tuned, the event-based rising edge detection system was able to correctly identify all target features in the MG time-series. The results in Figure 5.1.3 therefore demonstrate numerically that a single multi-modal RTD spiking neuron can be used to successfully implement a rising edge time-series

detection system.

This task was also experimentally demonstrated using an optically-sensitive RTD operating at the 1550nm telecom wavelength. In the experiments, the RTD was biased with a constant voltage of 0.61 V in the so-called IV peak region. We used a continuous background optical amplitude, with fast drops in optical power to trigger spike-firing events [114]. Similarly, the bias modulation was flipped such that a reduction in the bias voltage (moving further from the NDR region) increased the spiking firing threshold and optical amplitude requirement. The raw Mackey-Glass time-series values were encoded in the optical input and the delayed-copy of the signal in the electronic bias modulation. Similar to the numerically simulated results, 2 ns-long (negative) square pulses were used every 200 ns to encode the Mackey-Glass time-series value. An average optical input power of 2 mW was used. The electrical modulation varied the RTD with up to -0.3 V, creating a bias voltage range of 0.61 V to 0.31 V, large enough to inhibit even the largest optical inputs. Each RTD bias voltage was again held for 200 ns.

Figure 5.1.4 and 5.1.5 demonstrate the experimental achievement of the rising edge-feature detection task with the photodetecting RTD neuron. Figure 5.1.4(a) and (b) plot respectively, the raw values of the Mackey-Glass time-series signal and its delayed-copy, using a time-step of 200 ns. A visualisation of the delay difference is plotted in Figure 5.1.4(c). The first signal is optically-encoded into ‘negative’ optical pulses (2 ns per time-step) as shown in Figure 5.1.4(d). Here, higher values in the raw Mackey-Glass time-series signal are encoded with larger amplitude optical pulses. Figure 5.1.4(e) plots in turn the delayed values encoded electrically into the bias voltage modulation of the RTD, with higher time-series values encoded as lower voltages. Upon the simultaneous injection of the optical and electrical signals, the light-sensitive RTD neuron produced the temporal spiking output shown in Figure 5.1.4(f). The system revealed that excitable spikes were fired in response to the combinations of optical pulse powers and bias voltage inputs that represent a significant delayed difference in Mackey-Glass time-series signals. These points, that identify the rising edge-features in the time-series, occur when the amplitude of the optical input is large and the bias maintains operation near the peak point (closer to 0V modulation).

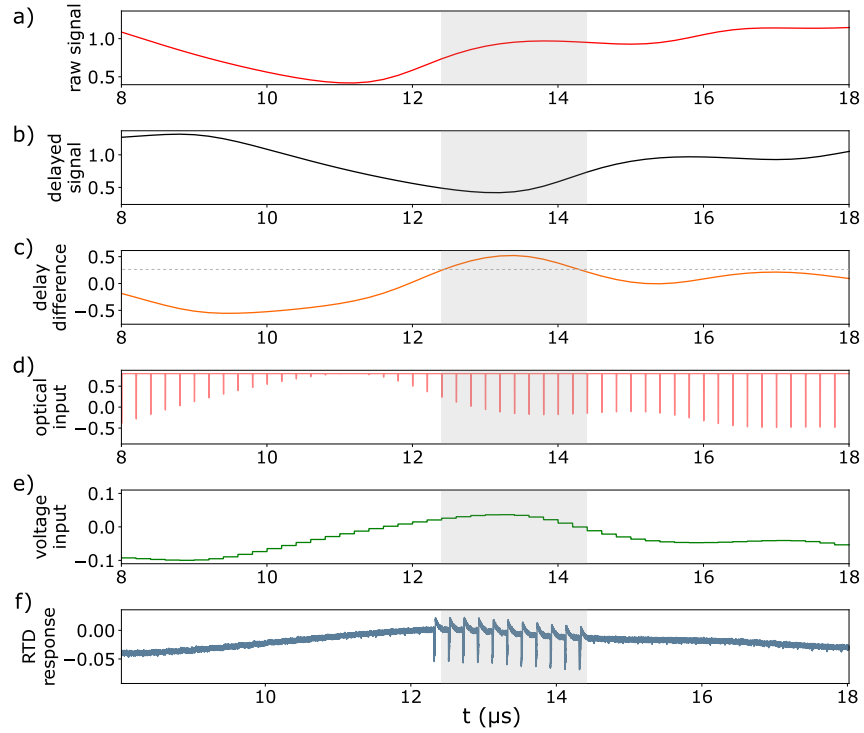


Figure 5.1.4: Experimentally measured RTD time-series showing rising edge-feature detection of Mackey-Glass time-series. a) Segment of raw MG time-series data. b) Delayed-copy of Mackey-Glass time-series (delayed by 10 time steps). c) Visualisation of the delayed difference between raw and delayed signals. d) The pulsed optical input used to encode the raw Mackey-Glass time-series. Pulse widths of 2 ns are separated by 200 ns. e) The modulation voltage used to encode the delayed signal into the bias voltage of the RTD. f) Measurement of the experimental (DC-filtered) voltage across the RTD, showing spikes at points when the time-series experiences a rising edge.

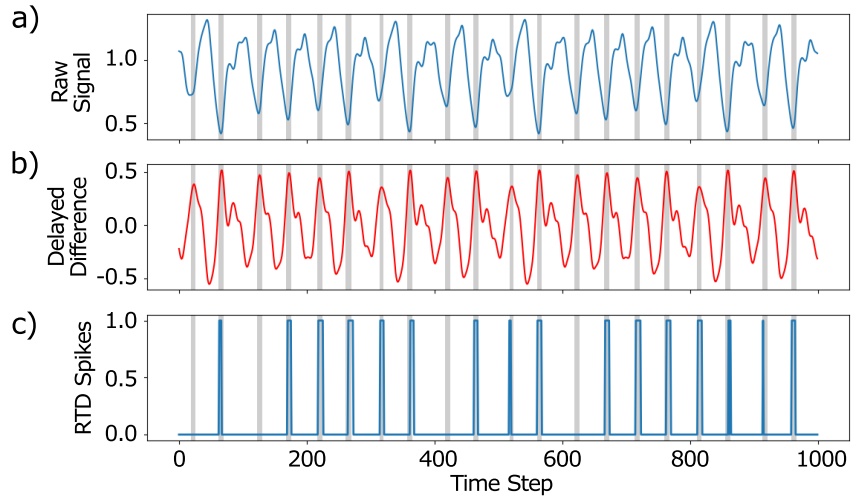


Figure 5.1.5: Rising edge detection of the Mackey-Glass time-series performed by an experimental RTD. a) The Mackey-Glass chaotic time-series signal. b) The calculated delayed difference between the raw MG signal and a copy delayed by 10 time steps. c) Time steps for which the experimental RTD emitted a spike ('0' represents no spike, '1' represents a spike), showing a proof of concept demonstration of dual modulation for rising edge detection. Highlighted in grey (a-c) are regions where detections are expected after exceeding a delayed difference of 0.32.

Overall, the experimental results of the task are summarised in Figure 5.1.5. Figure 5.1.5(a) shows the tested Mackey-Glass chaotic time-series, whilst Figure 5.1.5(b) plots the delayed difference between the raw and 10-time step delayed signals. Here, the grey shaded regions reveal time steps where rising edges (where the delayed difference exceeds 0.32), occur in the time-series. The time-series in Figure 5.1.5(c) reveals the time steps for which the RTD fired an excitable spike ('1' indicating spiking, '0' indicating non-spiking). Figure 5.1.5 reveals that the proof-of-concept experimental demonstration of the system was successful, having detected the majority of the rising edges in the time-series (16 out of 20 in this experimental proof-of-concept), without false positive detections. Further, due to the chaotic, slow-varying, nature of the time-series, we would expect to see improvements in performance when applied to real-world time-series that experience larger and more dramatic amplitude variations in short-temporal windows.

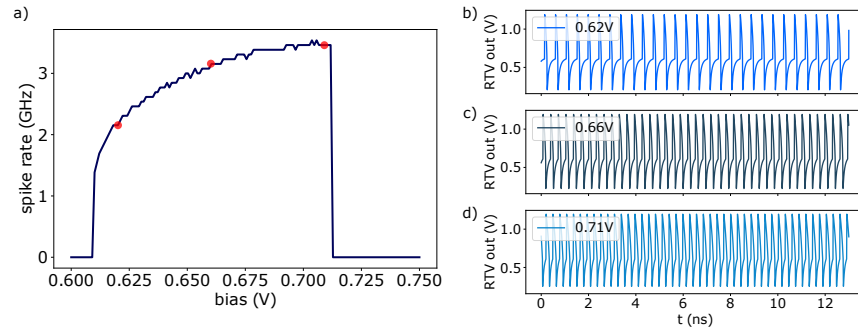


Figure 5.1.6: a) Spike rate against bias voltage. When the RTD is biased inside the NDR, it spikes at a constant rate as the circuit follows the limit cycle around the IV curve. As the voltage is increased inside the NDR, the spike frequency increases. b-d) Traces of the RTD voltage when biased at points inside the NDR (0.62V, 0.66V and 0.71V), marked by red points in a)

5.1.2 Spike Rate Encoding with RTD Neurons

Biological neurons often encode the information they are transmitting in the rate at which they emit spikes, a scheme known as spike-rate encoding[27, 127]. Rate encoding has also been used in biologically inspired spiking neural networks built for signal processing[128].

Here, an RTD neuron is set up as a rate encoding neuron, converting the value of the applied bias voltage into the frequency of emitted spikes. When the RTD is biased inside the region of negative resistance, the circuit undergoes relaxation oscillations (in which each period of the oscillation is a spike) with a constant frequency, shown in figure 5.1.6 (b-d). The spike rate (derived from the number of spikes measured in $10ns$) against bias voltage is shown in figure 5.1.6 (a), demonstrating the variance of the spike rate across $2GHz$ (from $1.4GHz$ at $0.61V$, to $3.4GHz$ at $0.7V$) as the bias voltage is swept through the NDR, enough range to facilitate rate encoding.

To demonstrate the rate coding scheme, a single RTD was used to encode a 50-symbol sequence of 8-level symbols. In order to do this, each of the 8 levels was mapped to a value of bias voltage inside the RTD's NDR (shown in figure 5.1.7 (a)), such that the RTD will spike at distinct rates for each possible symbol. Each symbol in the sequence was held for $5ns$, which given the RTD spike rate of several GHz, allows

for several spikes to occur, giving a good measure of the spike rate during the symbol.

The original signal can then be recovered from the time interval between successive spikes. The time of each spike is recorded, and the inverse of the interval since the previous spike is calculated, as an estimate of the rate of spikes. The decoded signal can then be formed by interpolating the pairs of spike times and inverse intervals.

Figure 5.1.7 shows the use of rate encoding and decoding of the 8-level sequence. Gaussian noise of different levels was also added to the signal, to test the robustness of the rate encoding scheme. The noise used had standard deviations of $\sigma = 0.05V$ (d,e) and $\sigma = 0.1V$ (f,g) (the underlying signal has a standard deviation of $\sigma = 0.06V$). Without any added noise, the decoding was able to almost perfectly recover the original signal (figure 5.1.7 (c)), save for overshoots at the transitions between different levels. With more and more noise, the reconstruction stays close to the true signal, but has more deviations. The eye diagrams in figure 5.1.7 (h) show how the reconstructed levels become more blurred as the noise level is increased, becoming almost indistinct with noise level $\sigma = 0.1V$. The eye diagrams also show the $1ns$ time needed for the spike rate to reliably transition between each level, setting the limit on how quickly the rate encoding can be modulated.

5.2 Neuromorphic Processing with Arrays of RTD Neurons

While single spiking neurons are capable of a number of computational tasks, larger networks consisting of several interconnected neurons are expected to have a wider range of applications and more complex processing capability. It is therefore desirable to demonstrate the operation of several spiking RTDs in parallel, in architectures where they are uncoupled (performing their own separate calculations simultaneously), and with coupling (optical or electrical) between devices forming larger spiking neural networks.

This section will demonstrate, through the numerical simulation described in Chapter 4, the use of arrays of several RTD neurons for two simple neuromorphic processing

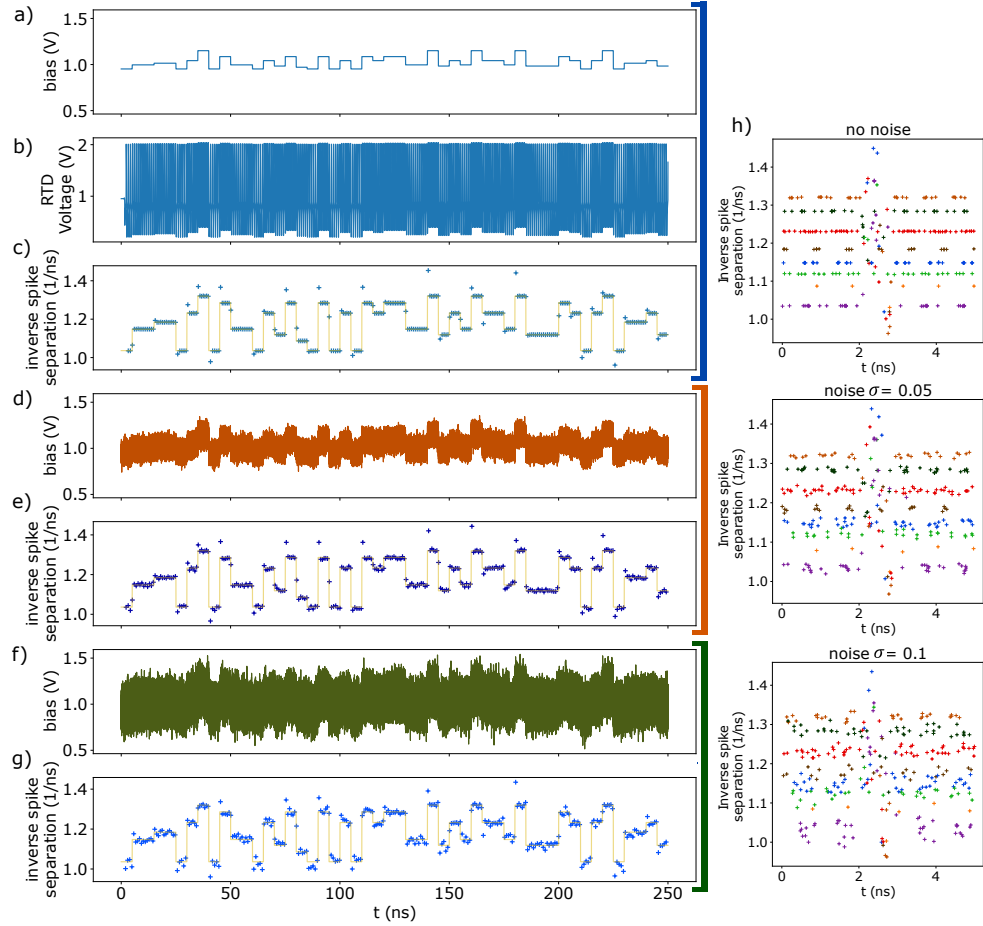


Figure 5.1.7: Simulation of a spike rate encoding of an 8 level word. a) Modulation of the bias voltage to the RTD. Each symbol was held for 5ns. b) Voltage across the RTD, showing the constant spiking at different rates, according to the bias modulation. c) Reconstruction of the signal from spike interval. At the time of occurrence of each spike, the reciprocal of the time since the previous spike is plotted, to approximate the inverse of the spike rate, recovering the modulation signal. d,f) Original modulation signal with normally distributed noise (standard deviation $\sigma = 0.05$, $\sigma = 0.1$) added. e,g) Reconstruction of signal by inverse spike interval, showing a good reconstruction of the non-noisy signal. h) corresponding eye diagrams. Gold lines in c,e,g show the target function to recover.

and memory tasks.

5.2.1 Photonic Spiking Neural Network with uncoupled RTD neuron arrays

This section explores the simplest possible use of multiple RTD neurons in parallel, in a system in which they are not connected to each other but instead operate individually simultaneously. This basic approach is used to build a photonic spiking neural network (SNN) able to perform a complex processing task. Specifically, we demonstrate numerically that a small array of only 20 uncoupled photodetecting RTD neurons can operate as a Photonic Spiking Extreme Learning Machine (ELM)[16] for a dataset classification task at ultrafast nanosecond rates. ELMs are simple neural network architectures, that share many key aspects to reservoir computers [129, 130], but differ by only enabling information to flow in a feed-forward manner (without recurrence) [16, 66]. As such, ELMs are composed of a single input layer (from which information enters the network), a hidden layer (with random untrained connections between nodes), and an output readout layer. In the ELM paradigm (as in reservoir computing) the input and hidden layer weights are random and fixed, only requiring the output layer weights to be trained. This is in stark contrast with common approaches based, for example, upon Deep Neural Networks (DNN), that require all connections to be trained, therefore requiring complex and costly learning procedures to calculate a very large number of weights. ELMs therefore dramatically reduce the network training requirements, whilst still allowing excellent performance in complex processing tasks (e.g. dataset classification) of widespread interest [66, 131].

Figure 5.2.1 shows the operation principle and architecture of the simulated RTD-based photonic spiking ELM proposed in this work. The schematic, shown in Figure 5.2.1(a), is highly hardware-friendly, with the input data injected optically into the network array of $N = 20$ uncoupled light sensitive RTD neurons. Each of the 20 RTD neurons simulated with equal properties and parameters, and all devices were biased with the same voltage level of 0.75 V (in the 'valley' region of their I-V relationship). The processing capability of the proposed photonic SNN was demonstrated using the

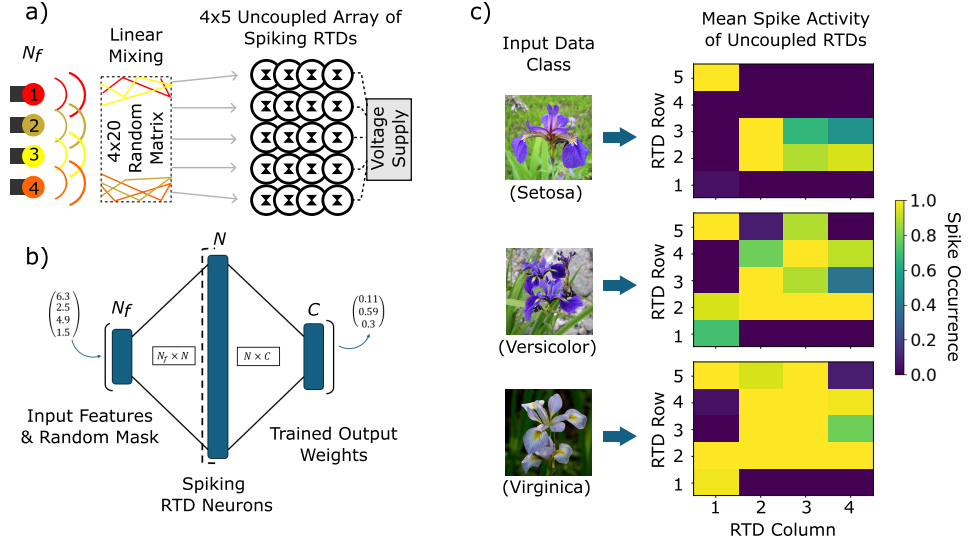


Figure 5.2.1: Schematic of the simulated RTD based neural network and analysis of RTD-array spikes for the Iris flower dataset classification task. a) Schematic of proposed operation. Iris flower input features (of which there are $N_f = 4$) are randomly scattered creating a linear mixing (weighting) of the inputs. Numerically, linearly mixed inputs are achieved through multiplication with a randomly generated matrix. The mixed inputs arrive optically on each of the ($N = 20$) RTD neurons in the array. b) The architecture of the proposed neural network, including feature mixing/masking (Input layer), RTD neuron array (hidden layer), and the trained readouts (output layer). c) Mean spike activity (occurrence) maps for each dataset class. Spike maps highlight differing responses from spatial RTDs within the array, as well as evolving patterns across dataset classes.

benchmark Iris flower dataset classification task [102], as done in previous sections. Figure 5.2.1(b) plots the architecture of the pSNN built with an array of uncoupled RTD neurons. The N_f data features are encoded using individual amplitude-modulated optical pulses, with a short temporal duration of just 50 ps . The optical input pulses, encoding the data features, are linearly mixed before entering the 20 uncoupled RTD neurons in the array. Experimentally, this can be achieved for example, by propagating the optical input pulses through a scattering medium prior to their arrival onto the RTD array. The linear data mixing stage is simulated by multiplying the data features (N_f) with a randomly generated weight matrix ($N_f \times N$) to produce a weighted input for each of the 20 RTDs. These weighted input values determine the amplitude of the optical pulses used to optically-trigger each spatial photo-detecting RTD in the

network. The weighted inputs are injected simultaneously into the photonic spiking ELM at a rate of 1 datapoint per ns, highlighting the high-speed (1 *GHz*) neuromorphic processing rate of this proposed system (the rate of data input could be increased further, by several GHz, as allowed by the ~ 300 ps refractory period of the RTD neuron [120]). Spatial spike patterns produced from the 20 RTDs are read-out directly from the simulated electrical output of each element in the array. The spike patterns are mapped to binary values ('1' for spiking, '0' for non-spiking) before an output layer weight, calculated following a simple supervised learning approach (linear least squares), is applied to readout the final classification of the datapoint.

Figure 5.2.1(c) plots the numerically-calculated mean spiking activity measured at the output of the RTD array, upon the injection of all 50 datapoints from each dataset class. The maps in Figure 5.2.1(c) show clear differences in the attained spiking patterns from the 20 RTDs for the three different dataset classes, clearly indicating the potential use of spiking activity for classification tasks. In particular, the spike firing and thresholding capabilities of RTD neurons are used as the activation function within the feed-forward neural network. Here, the hidden layer nodes are each represented by a single RTD neuron, with each one of them implementing the Heaviside step function to the amplitude of the input optical pulses:

$$H(x - T) = \begin{cases} 0 & x \leq T \\ 1 & T < x \end{cases} \quad (5.2)$$

where, if the input x is greater than the threshold T , the RTD will emit a spike. This is then used to construct a spatially multiplexed photonic spiking extreme learning machine (ELM), a feedforward neural network in which only the output layer weights are trained, while the input weights are fixed and selected at random. The small network size, created with a 4×5 array of RTDs, allows for reduced physical complexity and overall power usage, whilst still permitting operation at ultrafast speeds (GHz rates).

The numerically-calculated spiking output of the system, when tasked with the Iris flower dataset, is shown in more detail in Figure 5.2.2(a). Here, binary responses to the datapoints of each class are shown, revealing that different classes produce slightly different spiking patterns. These plots are produced by representing the spatial ELM as a 1×20 dimensional vector in which each element denotes whether the corresponding RTD neuron in the array spiked or not ('1' for spiking, '0' for non-spiking). Examples of the simulated RTD outputs for the cases of spiking and non-spiking (quiescent) responses are shown in Figure 5.2.2(b).

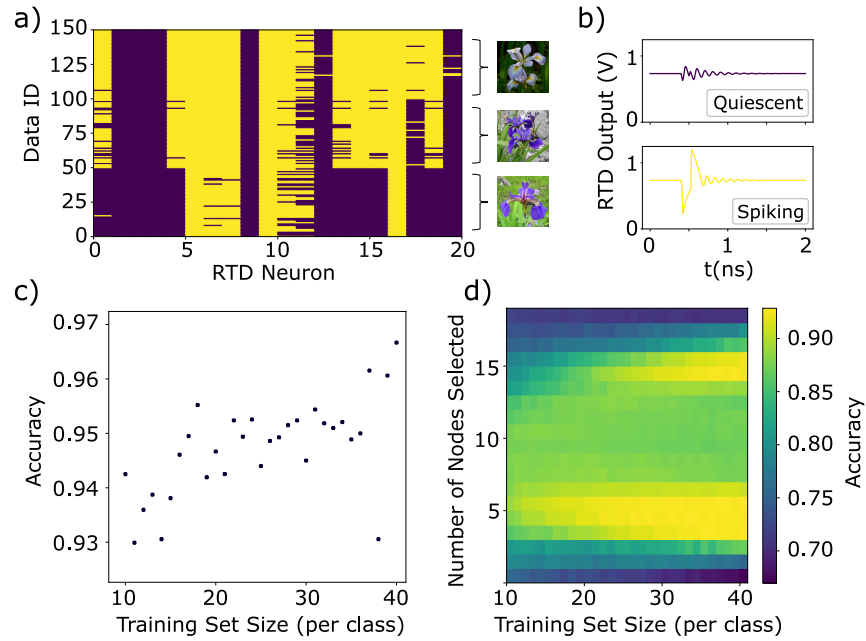


Figure 5.2.2: Iris flower dataset classification with a simulated array of RTDs. a) Spiking activity of the photonic spiking ELM generated by each of the 20 uncoupled RTD neurons in the array, for all 150 datapoints of the task (50 per class). Yellow indicates the detection of a spike-firing event, purple denotes no spike firing (quiescence), for the associated node and datapoint. (b) Example time-series of the simulated RTD neurons with spiking and non-spiking responses to datapoints. (c) Accuracy of the RTD ELM when trained with the least squares method. Maximum accuracy reached was 96.5%, for 39 points per class in the train set. d) Accuracy when trained with the node significance selection algorithm. Maximum accuracy reached was 93%, for 35 points per class for training and selecting 6 nodes only.

After gathering the binary output of the array, learning procedures were applied to train the output layer weights. The first training method used was applying a linear

least squares algorithm to the 1×20 vector to calculate an output weight matrix. The results achieved with the least squares training are plotted in Figure 5.2.2(c), showing the testing accuracy against the number of datapoints of each class used for the training dataset (with the remaining datapoints used for testing). The accuracy achieved with this approach, for the proposed photonic spiking ELM, was consistently above 93%, even for the cases of using small training set dimensions. Performance reached a peak of 96.5%, when trained with 39 points of each class, successfully demonstrating the computational potential of the proposed RTD-based photonic spiking ELM.

Furthermore, due to the discrete nature of the spiking responses from the photodetecting RTDs (either spiking or non-spiking), we can make use of the 'node significance' algorithm introduced previously in Section 3.2.3. This method introduces a new hyperparameter for the training method, the number of network nodes to select for the training phase, which can have a significant effect on performance. The map in Figure 5.2.2(d) shows the testing accuracy of the photonic spiking ELM when trained using this novel 'node significance' algorithm. Performance is measured against the number of points per class in the training set, and against the number of 'significant nodes' selected by the algorithm. The peak performance reached was 93% classification accuracy, for 35 points per class used for training and 6 nodes selected. This value matched the performance of the least squares training algorithm, however produced a more simple, sparse, binary-weight matrix. The increased sparsity, and binary output weights, have positive connotations for the implementation of these systems in hardware, as reduced training calculations and complexity can improve the energy efficiency and training time of the RTD array-based photonic SNN.

5.2.2 Neuromorphic Photonic Memory built with coupled RTD neurons

Memory in artificial neural networks is typically created by recurrent connections between neurons, in which the output of a neuron is fed back into itself (or, via a longer path of neurons that lead to the original). Recurrent neural networks have found applications in time-series processing task such as speech recognition and language pro-

cessing [132], however, they suffer from so called vanishing gradients that lead to slow training. The temporal nature of spiking neural networks on the other hand, leads to recurrent connections being handled naturally. A regenerative spiking memory system, composed of a single photonic RTD neuron with its optical output connected via delay line to its own optical input, was previously reported [133], constructing the simplest possible recurrent spiking neural network. In this way, spikes emitted by the photonic RTD neuron were able to repeatedly trigger new spikes after the delay has passed [133, 134]. This represents a memory system as the state (i.e. the existence and timing of the spike) retains information about the previous inputs to the photonic RTD neuron system.

Here, the single neuron autaptic memory, consisting of a photonic RTD connected to itself via an optical feedback loop, is recreated numerically. Figure 5.2.3(a) demonstrates the operation of this autaptic photonic spiking memory, with a delay (τ) of 5 ns, repeatedly storing a single spike event. The RTD neuron is initially triggered by an external optical pulse (dotted box in Figure 5.2.3(b)). The RTD neuron fires this first spike in response to a positive optical pulse of 10 ps duration. After this initial input, the system is only influenced by the delayed optical feedback, where a spike is regenerated by the previous cycle every 5 ns. Figure 5.2.3(c) maps the temporal RTD neuron output for each memory cycle. This map reveals that a slight delay of 131 ps occurs in each cycle due to the time taken for the RTD neuron to fire a spike. Spikes can therefore be successfully stored within an autaptic system created with an optical feedback loop and an RTD. Spiking feedback dynamics in similar photonic systems are explored in more detail in [134–136].

The limitation of the above single device memory, is that once a spike is introduced to the system, so long as there is no other perturbation, it will continue to be stored indefinitely. For information computing purposes it is preferable to have what is known as "fading memory" [137, 138] - where the effect of any input perturbation eventually fades away to zero, and the system has a single steady state. Here, a network of optically coupled photonic RTD neurons is demonstrated numerically, forming a regenerative memory cell that has a finite and tunable memory depth.

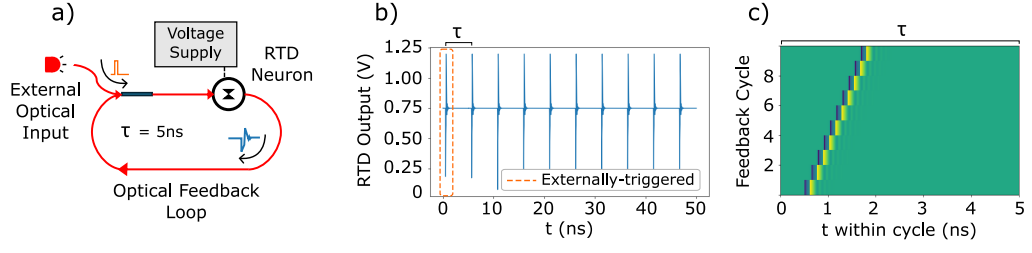


Figure 5.2.3: Autaptic spiking memory system with a simulated RTD neuron. a) Schematic of the simulated system. An external optical output is used to initially trigger a single spiking response. The RTD neuron output is fed back to itself via a 5 ns-long optical delay line. b) Simulated temporal output of the RTD neuron for 10 memory cycles, successfully regenerating the spike each time. c) Map of the temporal response of the RTD neuron for each cycle, showing an additional spike activation delay of 131 ps when the spike is regenerated.

Figure 5.2.4(a) shows a schematic diagram of the proposed system used to implement this adjustable neuromorphic photonic fading memory cell. Ten photonic RTD neurons are connected, all-to-all, through fan-in and fan-out optical connections (via optical fibres for example). An optical attenuator is implemented in the feedback loop such that the optical outputs are weighted equally and combined in a single optical path for reinjection. Each RTD neuron sees the same delayed optical input and the optical power landing on device equals the mean output of all devices (added to any external inputs). This system of 10 simulated photonic RTD neurons is referred to as a 'memory cell', as it is a self-contained structure that can exhibit photonic spiking memory. If all RTD neurons in the memory cell were the same (identical device features, given the same bias voltage and laser current), it would function identically to a single RTD in a self-feedback loop, where an excitation of one RTD would be continued in the loop indefinitely. For this reason, each device is driven with a different bias voltage. In combination with the attenuated optical feedback, this grants control over the temporal spiking responses of the RTD neurons in the cell, and importantly the overall spiking memory depth. Due to the bias voltage, each RTD neuron in the memory cell will be positioned further/closer from the boundary of the NDR, creating a range of spike activation threshold requirements for each neuron. This will cause specific subsets of the RTD neurons to elicit spikes when subject to higher/lower levels of

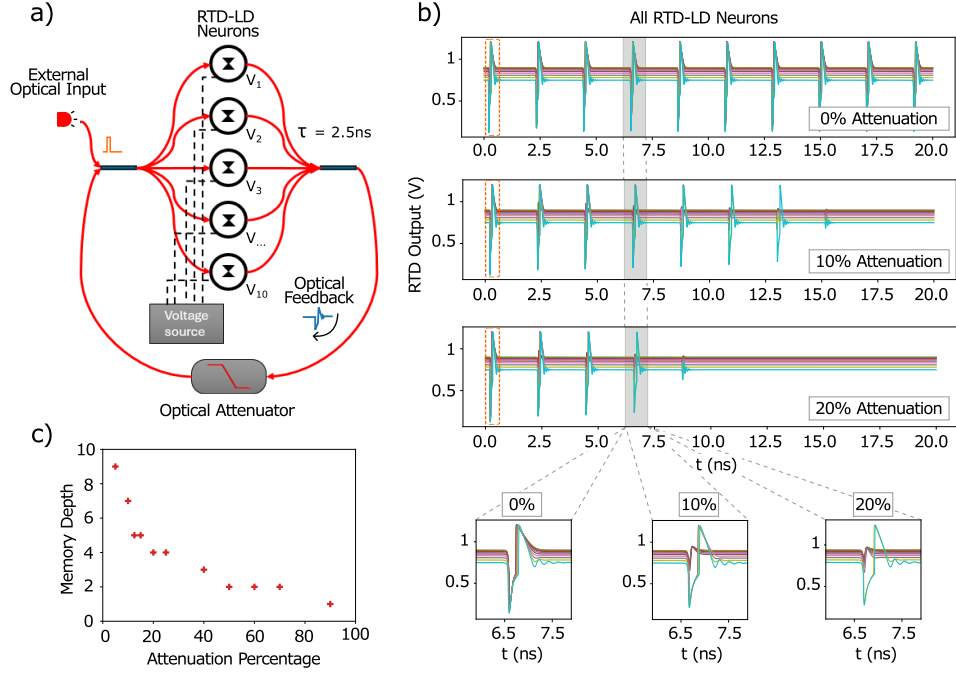


Figure 5.2.4: Neuromorphic fading memory cell constructed of 10 RTD neurons. (a) Schematic of the system with RTD neurons connected all-to-all via an attenuator. Each of the RTD neurons are biased at different voltages $V_{0n} = 0.9 - 0.00185n^2$, and experience a feedback loop of 2.5 ns. (b) Simulated RTD neuron outputs (10 devices overlaid) for 10 memory cycles. External optical input pulse is used to trigger the first response in each plot (dotted box). Top, middle and bottom plots (and respective insets) show results for increasing optical feedback attenuation (0, 10 and 20%). (c) Measured memory depth plotted against various feedback attenuation levels showing tunability.

optical feedback (i.e. more/fewer optical spikes stored in memory). Finally, depending on the strength of the optical attenuation, the overall number of devices successfully regenerating spikes into memory can be tuned, providing an effective neuromorphic photonic spiking fading memory cell system.

Figure 5.2.4(b) plots the numerically-calculated time-series demonstrating the operation of the 10 coupled RTD neuron photonic fading memory cell. Each device in the network was biased at progressively higher voltage, according to the formula $B_n = 0.9 - 0.00185n^2$, such that the lowest was set at 0.75 V (near the valley point), and the highest at 0.9 V (far from the NDR boundary with a high spike activation threshold). The parabolic distribution of the biases accounts for the nonlinear change

in spike threshold with bias. The system was set with a delayed feedback loop of 2.5 ns, and was subject to injection of a single 50 ps long optical input pulse at $t = 10$ ps. This positive optical pulse triggered the firing and storage of the first optical spike event in the network for every RTD neuron. Figure 5.2.4(b) shows the responses from each RTD neuron (overlaid) when the optical attenuation in the delayed feedback loop was increased from zero to 10% and 20%. Figure 5.2.4(b, top) reveals that all RTD neurons were triggered by the external optical input, and as a result of the high feedback strength, regenerated spikes every 2.5 ns. This shows that a permanent memory was achieved with a spike always regenerating each cycle, as in the single autaptic neuron case investigated before. For increased attenuation (Figure 5.2.4(b, middle) and 5.2.4(b, bottom)), the amplitude of the recurring optical spikes was reduced, therefore a different system memory depth becomes apparent. For the case of 10% attenuation in Figure 5.2.4(b, middle), the RTD neurons with the highest bias voltage did not receive enough optical power from the feedback loop to trigger a spike, even when all 10 neurons contributed to the initial input. The number of RTD neurons that fired therefore decreased with each delayed feedback cycle, enabling the fading optical memory. Figure 5.2.4(b, bottom) shows in turn that with even higher attenuation (20%), the number of firing RTD neurons decreased faster; hence reducing the total memory depth. In the fading memory cell feedback attenuation thus controls the 'depth' of the fading memory (how long a spike remains in the memory cell). For the case of the 10 device memory cell (under the investigated bias voltage values), further increasing the attenuation decreased the number of feedback cycles until spikes would no longer regenerate. Figure 5.2.4(c) plots the maximum number of cycles a spike was stored in memory against the optical attenuation strength. With zero attenuation (or any amplification of feedback), the memory depth is infinite, as the spike does not disappear at all. For attenuation over 90%, the spike is not regenerated at all, even in the first feedback cycle.

For the demonstrated system, the relationship between the memory depth and attenuation for the studied system was non-linear, likely as a result of the distribution of the biases for each RTD. Choosing the distribution of bias values would allow for

control of the attenuation - memory depth relation, and so more precise control of the memory depth.

5.3 Summary

This chapter demonstrated the potential of photonic-electronic RTD spiking neurons for neuromorphic information processing and memory systems. Using single RTD neurons, we showed the ability to perform fast rising edge detection in time-series data, both in simulation and experiment, with strong agreement between the two. We further establish spike rate encoding as a mechanism for information representation, and demonstrated the ability to encode and recover accurately a multi-level signal.

Beyond single device architectures, we also explored architectures of RTD neurons. Arrays of uncoupled devices were shown to implement a photonic spiking Extreme Learning Machine capable of high-speed dataset classification with excellent accuracy. Additionally, optically coupled RTD neurons were used to realise regenerative fading memory cells, and demonstrate how the memory depth and dynamics can be controlled through choice of bias voltages and optical attenuation.

Chapter 6

Conclusions

This final chapter provides a summary of the key results of this thesis, and discusses the potential lines of future research in neuromorphic photonics.

6.1 Conclusions

This thesis builds on previous research done by our group at Strathclyde[70, 71, 126, 139, 140], bringing the spiking photonic neurons based on VCSELs and Resonant Tunneling Diodes from proof-of-concept spiking neurons to functioning applications in photonic neural networks and neuromorphic systems.

Chapters 2 & 3 built on the previous work using VCSELs as ultrafast photonic spiking neurons at Strathclyde[69, 70, 101, 139, 141], applying the VCSEL-neurons as the nodes in photonic spiking neural networks. The first of these two chapters described the form and function of VCSELs, and experimentally characterised the photonic neuronal-like spiking responses: By injecting light into the cavity of the VCSEL at the right power level and detuning from the VCSEL's own emission modes, the VCSEL would be injection locked to the external signal. At the boundary of the conditions for this injection locking, input perturbations in the power of the injected light caused the VCSEL to undergo fast neuron-like spiking dynamics. The optical power spike emitted had a duration on the order of $\sim 100ps$, multiple orders of magnitude shorter than the spikes generated by biological neurons, and even those from electronic circuit-based

neurons. Experiments also showed that the spiking response of the VCSEL followed the same thresholding, integrate-and-fire and refractoriness behaviours featured by biological neurons. Thresholding means that the VCSEL neuron only emits a spike for inputs that are sufficiently strong, while all inputs less-than-threshold result in no output, while integrate-and-fire is the ability of the VCSEL to respond to the *integrated* value of a sequence of inputs perturbations, over a short amount of time. The integration time typical of these VCSEL neurons was found to be on the order of $300ps$. The refractory period - the shortest amount of time between two successive spikes - was shown to be $600ps$, allowing the VCSEL to spike at GHz rates.

These VCSEL neurons were then used in Chapter 3 to build fast photonic spiking neural networks based on the ELM paradigm, demonstrating artificial neural networks (ANNs) for classification and time-series prediction that could operate at GHz spiking rates. This demonstrated the experimental realisation of a VCSEL-based photonic spiking neural network (p-SNN) implemented using a single optically injected VCSEL and time-division multiplexed (TDM) virtual nodes, showing how a single semiconductor laser device, operated in its excitable spiking regime, can implement large-scale SNNs with node counts ranging from hundreds to several thousands, operating at sub-nanosecond virtual node times and GHz processing rates.

The p-SNN architecture is built upon that of time-division multiplexed reservoir computers, where the output of each node in the network is represented by the (binary) spiking state of the VCSEL output sampled at fixed temporal intervals. The input weights to this network are generated randomly, while the output weights are trained on the values of the virtual nodes and the training labels. Given the neuronal dynamics such as integration and refractoriness of the VCSEL occurring on timescales of approximately $1ns$, the virtual node timing used was $250ps$, forming strong temporal coupling between successive (virtual) nodes, forming wide network connectivity without any physical feedback loops. Networks with $N_v = 256$ to 4096 virtual spiking nodes were experimentally realised using identical hardware, with total data point processing times ranging from $64ns$ (for $N_v = 256$) to $1024ns$ (for $N_v = 4096$).

For nonlinear classification tasks, the system was first benchmarked using the Fisher

Iris dataset - a standard dataset that contains classes that cannot be separated by a linear discriminator, where networks with 512 and 1024 nodes achieved classification accuracies of 91.7% and 100%, respectively, using as little as 10 training samples per class. The results show that increasing the node count improves class separability by allowing a wider range of spiking patterns generated by the VCSEL spiking dynamics. High classification performance was achieved despite the binary nature of the output signals, the absence of recurrent optical feedback, and the relatively simple training methods, confirming that the nonlinear spiking dynamics alone are sufficient to implement complex decision boundaries.

The chapter further validated the scalability of the p-SNN architecture on the more challenging MADELON dataset, comprising 500 input features, most of which are non-informative. Using large node counts of 2048 and 4096, the p-SNN achieved peak classification accuracies of 94.5% and 95.7%, rivalling the best reported software-based benchmarks from the original NeurIPS 2003 challenge[104]. Notably, these results were obtained using very small training sets, in some cases as few as 1%–3% of the total dataset, highlighting the strong generalisation capability of the photonic spiking system.

A key development shown in this chapter is the introduction of a novel binary-weight training method tailored to the discrete spiking output of the VCSEL. Since the output of the VCSEL (a sequence of spikes) is discrete, and the desired output (the classification) is also discrete, it would be expected that the weights could also be chosen from a discrete set. This training method attempts to achieve that, by applying a modified chi-squared-based significance score to each node. The method identifies nodes whose spiking activity is most predictive of a given class, and sets the weight of those nodes to 1, leaving all others at 0. This approach enabled high-performance classification, setting the weights of fewer than 20 nodes per class to 1, corresponding to 0.5%–1% of the total available nodes, while maintaining or exceeding the performance of the standard ordinary least squares (OLS) training. This reduction in trained parameters significantly lowers computational cost, memory requirements, and training complexity, while producing sparse, binary output weight matrices that are well suited for future all-optical, hardware-based implementations.

Finally, the chapter extended the p-SNN framework beyond classification to time-series prediction, demonstrating multi-step-ahead prediction of the chaotic Mackey–Glass time-series. In spite of the lack of explicit optical feedback or memory in the feedforward p-SNN architecture, memory of the input data was effectively introduced through delayed-input encoding, where each input datapoint was combined with a delayed version of the time-series. Using this approach, networks with 256, 512, and 1024 nodes achieved strong predictive performance, with normalised mean squared errors (NMSE) as low as 0.051, and Pearson correlation coefficients up to 0.97 for 10-step-ahead predictions. These results are comparable to, and in some cases rival, traditional photonic reservoir computing systems that rely on optical delay lines and recurrent feedback.

Overall, this chapter has shown that VCSEL-based p-SNNs offer a highly efficient, scalable, and experimentally robust neuromorphic photonic computing platform. The system combines GHz-rate operation, sub-nanosecond temporal resolution, large virtual network sizes, and minimal training overhead, while maintaining strong performance across classification and prediction tasks. The demonstrated ability to trade off node count, processing time, training complexity, and accuracy provides a powerful degree of flexibility that is difficult to achieve in conventional electronic or photonic neural network implementations. These results firmly establish VCSEL-based p-SNNs as a viable and competitive hardware platform for ultrafast neuromorphic information processing and motivate further exploration toward integrated, fully photonic neuromorphic systems.

In Chapter 4, the Resonant Tunnelling Diode (RTD) was introduced. RTDs are devices formed by an integrated double-barrier quantum well (DBQW), that facilitates tunnelling of electrons across the barrier. The DBQW supports a bound electron state with an energy determined by the dimensions of the well. Electrons incident on the DBQW with energy near to that of the bound state have a high probability of tunnelling through the entire structure, with electrons that have energy exactly matching that of the bound state tunnelling with 100% probability - a phenomenon known as resonant tunnelling. This gives rise to a highly nonlinear current-voltage relation, including a region of negative differential resistance (NDR), in which the

current passing through the device decreases with increasing voltage - which occurs due to the decrease in resonant tunnelling as the bias voltage is moved away from the bound state energy. This NDR characteristic gives rise to a range of nonlinear dynamics such as relaxation oscillations and excitability. While the RTDs are purely electronic devices, they can be made to work with photonics by incorporating an optical window into the layered structure, allowing light to be injected into a photodetecting layer, injecting photogenerated carriers directly into the RTD.

RTD based spiking neurons were studied experimentally and in simulation, using a set of ordinary differential equations derived from a lumped-circuit treatment of the photonic-electronic setup, and a semi-classical derived model for the IV curve. This model was then used to demonstrate the full range of neuron-like behaviours that arise from biasing the RTD near its NDR range, using perturbations in the voltage and optical domain.

This chapter also demonstrated several core neuronal behaviours in RTDs: Thresholding was shown to arise from the basin of attraction surrounding the stable fixed point, with spike initiation requiring perturbations that cross a well-defined boundary in phase space. Maps of voltage and optical pulse amplitude versus duration revealed clear trade-offs, where shorter perturbations require higher amplitudes (either in the perturbation of bias or optical power) to induce spiking. Refractoriness was also demonstrated through paired pulse experiments, showing that spikes cannot be retriggered within a recovery window of approximately $300ps$, during which the system remains confined to the spiking trajectory.

The RTD neuron was further shown to operate as both leaky integrate-and-fire and resonate-and-fire neurons, depending on bias point. When biased at the IV peak, the RTD acted as an LIF neuron with respect to voltage perturbations, while when biased in the valley, the dynamics were more varied: At high bias voltages far from the NDR, overdamped oscillations allowed temporal integration of closely spaced sub-threshold inputs for operation as an LIF neuron, while operation deeper in the NDR valley produces underdamped oscillations with bias-dependent resonant frequencies, acting as an RaF neuron. In this regime, temporally aligned input perturbations can

constructively interfere to trigger spikes, enabling coincidence detection and frequency-selective responses.

Chapter 5 demonstrated several neuromorphic applications of the photonic-electronic spiking RTD neurons, based on simulated results and experimental data. These included event-based signal processing, construction of a spatially-multiplexed ANN, and a photonic spiking memory system. By exploiting the excitable dynamics of RTDs operating near the negative differential resistance (NDR) region of their IV curve, this chapter shows that both single-neuron functionality and larger spiking neural networks can be realised at GHz processing speeds.

In the same manner as with VCSELs, an ELM was constructed using RTD neurons as the activation functions. This network however was spatially multiplexed (instead of time-multiplexed), with each node in the network represented by a separate RTD. Since each RTD in the network ran in parallel, this network could be used for classification at GHz rates - as quickly as spikes could be triggered. This was tested against the Iris flower dataset, and could achieve upwards of 95% accuracy, using as few as 20 RTD neurons and training using ordinary least squares, and the novel significance scoring method.

One key feature of the RTD neurons dynamics was the ability to control the threshold of the amplitude of an optical pulse required to trigger a spike, by adjusting the bias voltage further or nearer to the NDR, in the IV curve valley. This was used initially to operate the RTD as a spiking edge detector, emitting a spike only when input time-series data underwent a large enough change over a short time. Due to the sub-nanosecond refractory period of the RTD neurons, this could process time-series data at gigasample-per-second rates, reliably detecting edge changes with a tunable threshold. The edge detection system was also implemented experimentally, using the dual optical and electrical modulation capability, running at MHz rates.

Further to this, the bias-variable threshold was used to create an optical feedback spiking memory system, with a tunable memory depth. In this system, several RTD neurons were connected all-to-all via a single optical channel that included an optical attenuator. Spikes emitted by lasers driven by the RTDs could then re-trigger further

spikes due to the feedback. The number of feedback cycles that a spike resulting from a single input could be re-generated was then tunable by adjusting the attenuation in the optical feedback line. This system is a more complex and powerful version of previously reported spiking regenerative memory, using only a single RTD with optical self-feedback, which would re-trigger a spike indefinitely.

Additionally, encoding of data using the spiking rate of the RTD was demonstrated, by biasing the RTD neuron inside its NDR. In this state, the RTD fired spikes continuously, at a rate determined by the bias position in the NDR, with the spike rate controllable from $1.4GHz$ up to $3.4GHz$ spiking rates. The rate encoding system was demonstrated in simulation on an 8-level word, running at $1ns$ per symbol (effectively 3GBit/s), with reliable decoding even with high noise ratios in the input voltage.

6.2 Future Work

The works presented in this thesis show the strong potential for the use of VCSELs and RTDs in the development of neuromorphic processing platforms, and leaves open a number of interesting routes for further development. Particularly, future work will likely focus on the applications of these devices to neuromorphics, such as construction of sensing or processing systems formed of larger networks of linked devices with less focus on new explorations of the dynamics of individual devices (which have been well-characterised).

6.2.1 Experimental and Theoretical Developments

In general, regardless of the specific device used to build spiking neurons, there is a large open question of the design and architecture of spiking neural networks, regarding how best to treat the inputs and outputs of an SNN and subsequently the processing methods and training algorithms. Conventional learning methods developed for artificial neural networks, such as backpropagation, rely on continuous valued activation functions that are differentiable, and precise weight updates. These assumptions are fundamentally at odds with the discrete, event-based nature of spikes, and often with

the constraints of photonic devices, meaning new training methods must be developed. Promising directions include local learning rules (e.g. Hebbian learning), surrogate-gradient methods adapted for SNNs, and physical training that incorporates measured hardware responses.

In parallel to the development of the theory of neuromorphic SNNs, the capability to build the necessary hardware should also be expanded. A chip-scale SNN will require not only the spiking devices themselves but also a means of connecting each device in a flexible and controllable manner - while this may be a challenge for modern integrated photonic circuits, systems of wide fan-in and fan-out on chip waveguides have been demonstrated[142] that could be useful for photonic SNNs. Weighting elements - or control of any number of parameters in the network - provide a similar hurdle, with the additional problem of scaling in number rapidly with network size (in a network formed by all-to-all single-parameter connections between devices, the number of weights scales quadratically with the device number). While a number of photonic devices that can be used for weighting exist (as discussed in section 1.3), integration with neuromorphic devices such as RTDs or VCSELs remains undeveloped.

6.2.2 Potential Applications of Neuromorphic Photonics

With regards to future applications, photonic neuromorphic systems offer strong advantages for sensing, particularly in contexts involving large amounts of data, while requiring low latency and high energy efficiency. Event-based sensing paradigms, such as neuromorphic or event-based cameras operate naturally with SNNs, as they both use asynchronous, sparse spiking signals. This change-driven representation dramatically reduces throughput and power consumption compared to traditional frame-by-frame video processing. This is a promising avenue for the use of the photosensitive RTDs, as the camera pixels could be replaced by RTD neurons that capture video information directly into spike trains, removing the need for intermediate processing.

The properties of photonics are also particularly useful for edge computing, where data from sensors must be processed locally and quickly, with minimal hardware complexity and power consumption. Applications in autonomous vehicles, robotics, drones,

and distributed sensor networks stand to benefit from photonic neuromorphic sensors that can perform early-stage perception and decision-making directly at the edge, reducing reliance on cloud-based processing and enabling fast responses to dynamic environments.

Spiking photonic neurons could also be applied to signal-to-spike encoding, providing an interface between real-world analog signals and spiking neural networks. Photonic neurons can convert temporal or amplitude variations into spike trains through their nonlinear and excitable dynamics, such as spike-rate encoding (as demonstrated with RTD neurons), or even through threshold-and-fire spiking. This would be particularly useful for autonomous systems, where continuous streams of sensor data, such as optical, RF, or lidar signals must be processed in real time with minimal latency. By embedding signal-to-spike encoding directly at the sensor or front-end hardware level, photonic neuromorphic systems can act as low-power, ultra-low-latency preprocessors that feed spiking networks operating on local devices. This tight integration of sensing, encoding, and computation has the potential to significantly improve the scalability, energy efficiency, and responsiveness of future intelligent autonomous and edge-based systems.

References

- [1] A. M. Turing, “Computing machinery and intelligence,” in *Parsing the Turing Test*, Springer Netherlands, 2009, pp. 23–65. DOI: 10.1007/978-1-4020-6710-5_3.
- [2] L. F. Menabrea, *Sketch of the Analytical Engine invented by Charles Babbage, Esq.* eng. Salt Lake City, UT: Project Gutenberg, 2025.
- [3] G. Cybenko, “Approximation by superpositions of a sigmoidal function,” *Mathematics of control, signals, and systems*, vol. 2, no. 4, pp. 303–314, 1989.
- [4] M. Minsky, S. Papert, and L. Bottou, *Perceptrons : an introduction to computational geometry*. The MIT Press, 2017.
- [5] S. Linnainmaa, “Taylor expansion of the accumulated rounding error,” *BIT*, vol. 16, no. 2, pp. 146–160, 1976. DOI: 10.1007/BF01931367.
- [6] D. E. Rumelhart and J. L. McClelland, “Learning internal representations by error propagation,” in *Parallel Distributed Processing*, MIT Press, 1987.
- [7] D. Rumelhart, G. Hinton, and R. Williams, “Learning representations by back-propagating errors.,” *Nature*, vol. 323, pp. 533–536, 1986. DOI: 10.1038/323533a0.
- [8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, *Imagenet classification with deep convolutional neural networks*, 2017. DOI: 10.1145/3065386.
- [9] A. Vaswani et al., “Attention is all you need,” 2023. [Online]. Available: <https://arxiv.org/abs/1706.03762>.

References

- [10] J. J. Hopfield, “Neural networks and physical systems with emergent collective computational abilities,” *Proceedings of the National Academy of Sciences*, vol. 79, no. 8, pp. 2554–2558, 1982. DOI: 10.1073/pnas.79.8.2554.
- [11] J. J. Hopfield, “Neurons with graded response have collective computational properties like those of two-state neurons,” *Proceedings of the National Academy of Sciences*, vol. 81, no. 10, pp. 3088–3092, 1984. DOI: 10.1073/pnas.81.10.3088.
- [12] Accessed: Nov. 9, 2025. [Online]. Available: <https://www.nobelprize.org/prizes/physics/2024/press-release/>.
- [13] L. Appeltant et al., “Information processing using a single dynamical node as complex system,” *Nature communications*, vol. 2, no. 1, p. 468, 2011. DOI: 10.1038/ncomms1476.
- [14] G. Van der Sande, D. Brunner, and M. C. Soriano, “Advances in photonic reservoir computing,” *Nanophotonics (Berlin, Germany)*, vol. 6, no. 3, pp. 561–576, 2017. DOI: 10.1515/nanoph-2016-0132.
- [15] W. Maass and J. Principe, “Special issue on echo state networks and liquid state machines,” *Neural networks*, vol. 20, no. 3, pp. 287–289, 2007. DOI: 10.1016/j.neunet.2007.04.001.
- [16] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: Theory and applications,” *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006. DOI: 10.1016/j.neucom.2005.12.126.
- [17] G. Hayes, *Scottish data centres powering AI already using enough water to fill 27 million bottles a year*. [Online]. Available: <https://www.bbc.co.uk/news/articles/c77zxx43x4vo>.
- [18] S. B. Furber, F. Galluppi, S. Temple, and L. A. Plana, “The spinnaker project,” *Proceedings of the IEEE*, vol. 102, no. 5, pp. 652–665, 2014. DOI: 10.1109/JPROC.2014.2304638.

References

- [19] S. Höppner et al., *The spinnaker 2 processing element architecture for hybrid digital neuromorphic computing*, 2022. arXiv: 2103.08392. [Online]. Available: <https://arxiv.org/abs/2103.08392>.
- [20] C. Pehle et al., “The BrainScaleS-2 accelerated neuromorphic system with hybrid plasticity,” 2022. arXiv: 2201.11063. [Online]. Available: <http://arxiv.org/abs/2201.11063>.
- [21] M. V. DeBole et al., “Truenorth: Accelerating from zero to 64 million neurons in 10 years,” *Computer*, vol. 52, no. 5, pp. 20–29, 2019. DOI: 10.1109/MC.2019.2903009.
- [22] D. S. Modha et al., “Neural inference at the frontier of energy, space, and time,” *Science*, vol. 382, no. 6668, pp. 329–335, 2023. DOI: 10.1126/science.adh1174.
- [23] M. Davies et al., “Loihi: A neuromorphic manycore processor with on-chip learning,” *Ieee Micro*, vol. 38, no. 1, pp. 82–99, 2018.
- [24] G. Orchard et al., “Efficient Neuromorphic Signal Processing with Loihi 2,” in *2021 IEEE Workshop on Signal Processing Systems (SiPS)*, IEEE, 2021, pp. 254–259. DOI: 10.1109/SiPS52927.2021.00053.
- [25] Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo, and J. Zhang, “Edge intelligence: Paving the last mile of artificial intelligence with edge computing,” *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1738–1762, 2019. DOI: 10.1109/JPROC.2019.2918951.
- [26] W. Maass, “Networks of spiking neurons: The third generation of neural network models,” *Neural networks*, vol. 10, no. 9, pp. 1659–1671, 1997. DOI: 10.1016/S0893-6080(97)00011-7.
- [27] A. Kepecs and J. Lisman, “Information encoding and computation with spikes and bursts,” *Network (Bristol)*, vol. 14, no. 1, pp. 103–118, 2003. DOI: 10.1080/net.14.1.103.118.

References

- [28] W. Guo, M. E. Fouda, A. M. Eltawil, and K. N. Salama, “Efficient training of spiking neural networks with temporally-truncated local backpropagation through time,” *Frontiers in neuroscience*, vol. 17, p. 1047008, 2023. DOI: 10.3389/fnins.2023.1047008.
- [29] A. Mohemmed, S. Schliebs, S. Matsuda, and N. Kasabov, “Span: Spike pattern association neuron for learning spatio-temporal spike patterns,” *International journal of neural systems*, vol. 22, no. 4, pp. 1250012–1250012-17, 2012. DOI: 10.1142/S0129065712500128.
- [30] S. B. Shrestha and G. Orchard, “Slayer: Spike layer error reassignment in time,” *arXiv.org*, 2018.
- [31] D. Hebb, *The Organization of Behavior*. New York: Wiley and Sons, 1949.
- [32] T. Masquelier, E. Hugues, G. Deco, and S. J. Thorpe, “Oscillations, phase-of-firing coding, and spike timing-dependent plasticity: An efficient learning scheme,” *The Journal of neuroscience*, vol. 29, no. 43, pp. 13484–13493, 2009. DOI: 10.1523/JNEUROSCI.2207-09.2009.
- [33] W. van der Veen, “Including stdp to eligibility propagation in multi-layer recurrent spiking neural networks,” *arXiv.org*, 2022.
- [34] F. Ponulak and A. Kasiński, “Supervised learning in spiking neural networks with resume: Sequence learning, classification, and spike shifting,” *Neural Computation*, vol. 22, no. 2, pp. 467–510, 2010. DOI: 10.1162/neco.2009.11-08-901.
- [35] A. L. Hodgkin and A. F. Huxley, “A quantitative description of membrane current and its application to conduction and excitation in nerve,” *The Journal of physiology*, vol. 117, no. 4, pp. 500–544, 1952. DOI: 10.1113/jphysiol.1952.sp004764.
- [36] R. FitzHugh, “Impulses and physiological states in theoretical models of nerve membrane,” *Biophysical journal*, vol. 1, no. 6, pp. 445–466, 1961. DOI: 10.1016/S0006-3495(61)86902-6.

References

- [37] D. Brunner et al., *Roadmap on neuromorphic photonics*, 2025. arXiv: 2501.07917. [Online]. Available: <https://arxiv.org/abs/2501.07917>.
- [38] H. J. Caulfield and S. Dolev, “Why future supercomputing requires optics,” *Nature photonics*, vol. 4, no. 5, pp. 261–263, 2010. DOI: 10.1038/nphoton.2010.94.
- [39] R. Osgood and X. Meng, *Principles of photonic integrated circuits : materials, device physics, guided wave design* (Graduate Texts in Physics), 1st ed. 2021. Springer, 2022. DOI: 10.1007/978-3-030-65193-0.
- [40] J. Feldmann et al., “Parallel convolutional processing using an integrated photonic tensor core,” *Nature (London)*, vol. 589, pp. 52–58, 2021. DOI: 10.1038/s41586-020-03070-1.
- [41] P. R. Prucnal, B. J. Shastri, T. Ferreira de Lima, M. A. Nahmias, and A. N. Tait, “Recent progress in semiconductor excitable lasers for photonic spike processing,” *Advances in optics and photonics*, vol. 8, no. 2, pp. 228–299, 2016. DOI: 10.1364/AOP.8.000228.
- [42] W. Coomans, L. Gelens, S. Beri, J. Danckaert, and G. Van der Sande, “Solitary and coupled semiconductor ring lasers as optical spiking neurons,” *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 84, no. 3, p. 036 209, 2011. DOI: 10.1103/PhysRevE.84.036209.
- [43] K. Alexander, T. Van Vaerenbergh, M. Fiers, P. Mechet, J. Dambre, and P. Bienstman, “Excitability in optically injected microdisk lasers with phase controlled excitatory and inhibitory response,” *Optics express*, vol. 21, no. 22, pp. 26 182–26 191, 2013. DOI: 10.1364/OE.21.026182.
- [44] A. Skalli et al., “Photonic neuromorphic computing using vertical cavity semiconductor lasers,” *Optical materials express*, vol. 12, no. 6, pp. 2395–2414, 2022. DOI: 10.1364/OME.450926.
- [45] M. Pflüger, D. Brunner, T. Heuser, J. A. Lott, S. Reitzenstein, and I. Fischer, “Experimental reservoir computing with diffractively coupled vcsels,” *Optics letters*, vol. 49, no. 9, pp. 2285–2288, 2024.

References

- [46] J. Bueno, J. Robertson, M. Hejda, and A. Hurtado, “Comprehensive performance analysis of a vcsel-based photonic reservoir computer,” *IEEE Photonics Technology Letters*, vol. 33, no. 16, pp. 920–923, 2021. DOI: 10.1109/LPT.2021.3075095.
- [47] A. Skalli, X. Porte, N. Haghighi, S. Reitzenstein, J. A. Lott, and D. Brunner, “Computational metrics and parameters of an injection-locked large area semiconductor laser for neural network computing,” *Optical materials express*, vol. 12, no. 7, pp. 2793–2804, 2022. DOI: 10.1364/OME.451524.
- [48] X. Porte, A. Skalli, N. Haghighi, S. Reitzenstein, J. A. Lott, and D. Brunner, “A complete, parallel and autonomous photonic neural network in a semiconductor multimode laser,” *JPhys photonics*, vol. 3, no. 2, p. 24 017, 2021. DOI: 10.1088/2515-7647/abf6bd.
- [49] K. Harkhoe, G. Verschaffelt, and G. Van der Sande, “Neuro-inspired computing with spin-vcsels,” *Applied sciences*, vol. 11, no. 9, p. 4232, 2021. DOI: 10.3390/app11094232.
- [50] M. San Miguel, Q. Feng, and J. V. Moloney, “Light-polarization dynamics in surface-emitting semiconductor lasers,” *Physical review. A, Atomic, molecular, and optical physics*, vol. 52, no. 2, pp. 1728–1739, 1995. DOI: 10.1103/PhysRevA.52.1728.
- [51] M. Hejda et al., “Optical spike amplitude weighting and neuromimetic rate coding using a joint vcsel-mrr neuromorphic photonic system,” *Neuromorphic computing and engineering*, vol. 4, no. 2, p. 24 011, 2024. DOI: 10.1088/2634-4386/ad4b5b.
- [52] A. N. Tait et al., “Neuromorphic photonic networks using silicon photonic weight banks,” *Scientific reports*, vol. 7, no. 1, pp. 7430–10, 2017. DOI: 10.1038/s41598-017-07754-z.
- [53] S. Biasi, G. Donati, A. Lugnan, M. Mancinelli, E. Staffoli, and L. Pavesi, “Photonic neural networks based on integrated silicon microresonators,” *Intelligent computing*, vol. 3, 2024. DOI: 10.34133/icomputing.0067.

References

- [54] G. Donati, C. R. Mirasso, M. Mancinelli, L. Pavesi, and A. Argyris, “Microring resonators with external optical feedback for time delay reservoir computing,” *Optics express*, vol. 30, no. 1, pp. 522–537, 2022. DOI: 10.1364/OE.444063.
- [55] A. Jha, C. Huang, H.-T. Peng, B. Shastri, and P. R. Prucnal, “Photonic spiking neural networks and graphene-on-silicon spiking neurons,” *Journal of light-wave technology*, vol. 40, no. 9, pp. 2901–2914, 2022. DOI: 10.1109/JLT.2022.3146157.
- [56] G. Donati, S. Biasi, L. Pavesi, and A. Hurtado, “All-optical spiking processing and reservoir computing with a passive silicon microring and wavelength-time division multiplexing,” *Photonics research (Washington, DC)*, vol. 13, no. 9, pp. 2641–2653, 2025. DOI: 10.1364/PRJ.558405.
- [57] M. Hejda et al., *Sephia: j1 laser/neuron spiking electro-photonic integrated multi-tiled architecture for scalable optical neuromorphic computing*, 2025. arXiv: 2510.07427 [cs.ET]. [Online]. Available: <https://arxiv.org/abs/2510.07427>.
- [58] D. Pierangeli, G. Marcucci, D. Brunner, and C. Conti, “Noise-enhanced spatial-photonics ising machine,” *Nanophotonics (Berlin, Germany)*, vol. 9, no. 13, pp. 4109–4116, 2020. DOI: 10.1515/nanoph-2020-0119.
- [59] R. Talukder, A. Skalli, X. Porte, S. Thorpe, and D. Brunner, “A spiking photonic neural network of 40.000 neurons, trained with rank-order coding for leveraging sparsity,” *Neuromorph. Comput. Eng.*, vol. 5, p. 034 003, 2025.
- [60] A. Skalli, M. Goldmann, N. Haghighi, S. Reitzenstein, J. A. Lott, and D. Brunner, “Annealing-inspired training of an optical neural network with ternary weights,” *Communications physics*, vol. 8, no. 1, pp. 68–10, 2025. DOI: 10.1038/s42005-025-01972-y.
- [61] Y. Shen et al., “Deep learning with coherent nanophotonic circuits,” *Nature photonics*, vol. 11, no. 7, pp. 441–446, 2017. DOI: 10.1038/nphoton.2017.93.
- [62] Z. Fan, J. Lin, J. Dai, T. Zhang, and K. Xu, “Photonic hopfield neural network for the ising problem,” *Optics express*, vol. 31, no. 13, pp. 21 340–21 350, 2023. DOI: 10.1364/OE.491554.

References

- [63] S. Pai et al., “Experimentally realized in situ backpropagation for deep learning in photonic neural networks,” *Science (American Association for the Advancement of Science)*, vol. 380, no. 6643, pp. 398–404, 2023. DOI: 10.1126/science.ade8450.
- [64] F. Böhm, G. Verschaffelt, and G. Van der Sande, “A poor man’s coherent ising machine based on opto-electronic feedback systems for solving optimization problems,” *Nature communications*, vol. 10, no. 1, pp. 3538–9, 2019. DOI: 10.1038/s41467-019-11484-3.
- [65] A. Lupo, E. Picco, M. Zajnulina, and S. Massar, “Deep photonic reservoir computer based on frequency multiplexing with fully analog connection between layers,” *Optica*, vol. 10, no. 11, pp. 1478–1485, 2023. DOI: 10.1364/OPTICA.489501.
- [66] D. Owen-Newns, J. Robertson, M. Hejda, and A. Hurtado, “Ghz rate neuromorphic photonic spiking neural network with a single vertical-cavity surface-emitting laser (vcsel),” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 29, no. 2: Optical Computing, pp. 1–10, 2023. DOI: 10.1109/JSTQE.2022.3205716.
- [67] D. Owen-Newns, J. Robertson, M. Hejda, and A. Hurtado, “Photonic spiking neural networks with highly efficient training protocols for ultrafast neuromorphic computing systems,” *Intelligent Computing*, vol. 2, p. 0031, 2023. DOI: 10.34133/icomputing.0031.
- [68] D. Owen-Newns et al., “Photonic spiking neural network built with a single vcsel for high-speed time series prediction,” *Communications physics*, vol. 8, no. 1, pp. 110–9, 2025.
- [69] J. Robertson, M. Hejda, J. Bueno, and A. Hurtado, “Ultrafast optical integration and pattern classification for neuromorphic photonics based on spiking VCSEL neurons,” *Scientific Reports*, vol. 10, no. 1, p. 6098, 2020. DOI: 10.1038/s41598-020-62945-5.

References

- [70] J. Robertson et al., “Ultrafast neuromorphic photonic image processing with a vcsel neuron,” *Scientific reports*, vol. 12, no. 1, pp. 4874–10, 2022. DOI: 10.1038/s41598-022-08703-1.
- [71] M. j. Hejda, J. Robertson, J. Bueno, and A. Hurtado, “Spike-based information encoding in vertical cavity surface emitting lasers for neuromorphic photonic systems,” *JPhys photonics*, vol. 2, no. 4, p. 44 001, 2020. DOI: 10.1088/2515-7647/aba670.
- [72] H. Soda, K.-i. Iga, C. Kitahara, and Y. Suematsu, “GaInAsP/InP surface emitting injection lasers,” *Japanese Journal of Applied Physics*, vol. 18, no. 12, pp. 2329–2330, 1979. DOI: 10.1143/jjap.18.2329.
- [73] G. Pan, M. Xun, X. Zhou, Y. Sun, Y. Dong, and D. Wu, “Harnessing the capabilities of vcsels: Unlocking the potential for advanced integrated photonic devices and systems,” *Light, science & applications*, vol. 13, no. 1, pp. 229–26, 2024. DOI: 10.1038/s41377-024-01561-8.
- [74] H. Soda, Y. Motegi, and K. Iga, “GaInAsP/InP surface emitting injection lasers with short cavity length,” *IEEE journal of quantum electronics*, vol. 19, no. 6, pp. 1035–1041, 1983.
- [75] W. Hofmann, M. Muller, G. Bohm, M. Ortsiefer, and M.-C. Amann, “1.55- μ m vcsel with enhanced modulation bandwidth and temperature range,” *IEEE photonics technology letters*, vol. 21, no. 13, pp. 923–925, 2009. DOI: 10.1109/LPT.2009.2020605.
- [76] A. Rouxel, K. B. Saddik, S. Calvez, A. Monmayrant, and G. Almuneau, “High-speed tunable interferometric spectral shaper for vcsel manufacturing inspection,” in *Conference on Lasers & Electro-Optics Europe & International Quantum Electronics Conference (Online)*, IEEE, 2025, pp. 1–1. DOI: 10.1109/CLEO/Europe-EQEC65582.2025.11111359.
- [77] E. Haglund et al., “30 ghz bandwidth 850 nm vcsel with sub-100 fj/bit energy dissipation at 25–50 gbit/s,” *Electronics letters*, vol. 51, no. 14, pp. 1096–1098, 2015. DOI: 10.1049/el.2015.0785.

References

- [78] T. Heuser, M. Pflüger, I. Fischer, J. A. Lott, D. Brunner, and S. Reitzenstein, “Developing a photonic hardware platform for brain-inspired computing based on 5×5 vcsel arrays,” *JPhys photonics*, vol. 2, no. 4, p. 44 002, 2020. DOI: 10.1088/2515-7647/aba671.
- [79] J. Wu, Y. Zeng, P. Zhou, and N. Li, “Broadband chaos generation in vcsels with intensity-modulated optical injection,” *Optics and laser technology*, vol. 159, p. 108 994, 2023. DOI: 10.1016/j.optlastec.2022.108994.
- [80] S. S. Alharthi et al., “Control of emitted light polarization in a 1310nm dilute nitride spin-vertical cavity surface emitting laser subject to circularly polarized optical injection,” *Applied physics letters*, vol. 105, no. 18, 2014. DOI: 10.1063/1.4901192.
- [81] A. Hurtado, K. Schires, I. D. Henning, and M. J. Adams, “Investigation of vertical cavity surface emitting laser dynamics for neuromorphic photonic systems,” *Applied physics letters*, vol. 100, no. 10, pp. 103 703–103703-4, 2012. DOI: 10.1063/1.3692726.
- [82] K. Iga, H. Soda, T. Terakado, and S. Shimizu, “Lasing characteristics of improved gainasp/inp surface emitting injection lasers,” *Electronics letters*, vol. 19, no. 13, pp. 457–458, 1983. DOI: 10.1049/e1:19830311.
- [83] B. Dayal Padullaparthi, J. Tatum, and K. Iga, “Automotive lidars,” in *VCSEL Industry*, ser. The ComSoc Guides to Communications Technologies, 1st ed., Wiley, 2021.
- [84] C. Wu et al., “Addressable structured light system using metasurface optics and an individually addressable vcsel array,” *Photonics research (Washington, DC)*, vol. 12, no. 6, pp. 1129–1137, 2024. DOI: 10.1364/PRJ.516942.
- [85] G. M. Isoe, S. Wassin, R. R. G. Gamatham, A. W. R. Leitch, and T. B. Gibbon, “A high capacity data centre network: Simultaneous 4-pam data at 20 gbps and 2 ghz phase modulated rf clock signal over a single vcsel carrier,” *Journal of modern optics*, vol. 64, no. 21, pp. 2336–2344, 2017. DOI: 10.1080/09500340.2017.1357855.

References

- [86] B. D. Padullaparthi, *VCSEL industry : communication and sensing* (ComSoc guides to communications technologies). Wiley : IEEE Press, 2022.
- [87] A. Hurtado, I. D. Henning, and M. J. Adams, “Optical neuron using polarisation switching in a 1550nm-vcsel,” *Optics express*, vol. 18, no. 24, pp. 25 170–25 176, 2010. DOI: 10.1364/OE.18.025170.
- [88] M. O. Markus Ortsiefer, R. S. Robert Shau, G. B. Gerhard Böhm, F. K. Fabian Köhler, G. A. Gerhard Abstreiter, and M.-C. A. Markus-Christian Amann, “Low-resistance inga(al)as tunnel junctions for long wavelength vertical-cavity surface-emitting lasers,” *Japanese Journal of Applied Physics*, vol. 39, no. 4R, pp. 1727–1729, 2000. DOI: 10.1143/JJAP.39.1727.
- [89] K. H. Rhew, S. C. Jeon, O. K. Kwon, D. H. Lee, B. S. Too, and I. Yun, “Reliability assessment of 1.55- μ m vertical cavity surface emitting lasers for optical communication systems,” in *2007 IEEE International Reliability Physics Symposium Proceedings, 45th Annual*, ser. Annual Proceedings - Reliability Physics (Symposium), 45th Annual IEEE International Reliability Physics Symposium 2007, IRPS ; Conference date: 15-04-2007 Through 19-04-2007, 2007, pp. 476–479. DOI: 10.1109/RELPHY.2007.369937.
- [90] B. Hong and A. Hajimiri, “A general theory of injection locking and pulling in electrical oscillators-part i: Time-synchronous modeling and injection waveform design,” *IEEE journal of solid-state circuits*, vol. 54, no. 8, pp. 2109–2121, 2019. DOI: 10.1109/JSSC.2019.2908753.
- [91] C.-H. Chang, L. Chrostowski, and C. Chang-Hasnain, “Injection locking of vcsels,” *IEEE journal of selected topics in quantum electronics*, vol. 9, no. 5, pp. 1386–1393, 2003. DOI: 10.1109/JSTQE.2003.819510.
- [92] C. Henry, “Theory of the linewidth of semiconductor lasers,” *IEEE Journal of Quantum Electronics*, vol. 18, no. 2, pp. 259–264, 1982. DOI: 10.1109/JQE.1982.1071522.

References

- [93] F. Mogensen, H. Olesen, and G. Jacobsen, “Locking conditions and stability properties for a semiconductor laser with external light injection,” *IEEE journal of quantum electronics*, vol. 21, no. 7, pp. 784–793, 1985. DOI: 10.1109/JQE.1985.1072760.
- [94] Z. G. Pan et al., “Optical injection induced polarization bistability in vertical-cavity surface-emitting lasers,” *Applied physics letters*, vol. 63, no. 22, pp. 2999–3001, 1993. DOI: 10.1063/1.110264.
- [95] J. Ohtsubo, *Semiconductor Lasers: Stability, Instability and Chaos*. Springer Cham, 2017. DOI: 10.1007/978-3-319-56138-7.
- [96] M. Ghisoni, J. Halonen, K. Shore, A. Larsson, and Y. Hong, “Polarisation switching in a vertical cavity surface emitting semiconductor laser by frequency detuning,” *IEE proceedings. Optoelectronics*, vol. 148, no. 1, pp. 31–34, 2001. DOI: 10.1049/ip-opt:20010246.
- [97] A. Hurtado, I. Henning, and M. Adams, “Two-wavelength switching with a 1550 nm vcsel under single orthogonal optical injection,” *IEEE journal of selected topics in quantum electronics*, vol. 14, no. 3, pp. 911–917, 2008. DOI: 10.1109/JSTQE.2008.918312.
- [98] A. Hurtado and J. Javaloyes, “Controllable spiking patterns in long-wavelength vertical cavity surface emitting lasers for neuromorphic photonics systems,” *Applied physics letters*, vol. 107, no. 24, 2015. DOI: 10.1063/1.4937730.
- [99] C. Teeter et al., “Generalized leaky integrate-and-fire models classify multiple neuron types,” *Nature communications*, vol. 9, no. 1, pp. 709–15, 2018. DOI: 10.1038/s41467-017-02717-4.
- [100] J. A. Alanis, J. Robertson, M. Hejda, and A. Hurtado, “Weight adjustable photonic synapse by nonlinear gain in a vertical cavity semiconductor optical amplifier,” *Applied physics letters*, vol. 119, no. 20, 2021. DOI: 10.1063/5.0064374.
- [101] J. Robertson, E. Wade, Y. Kopp, J. Bueno, and A. Hurtado, “Towards Neuromorphic Photonic Networks of Ultrafast Spiking Laser Neurons,” *IEEE Journal*

References

- of Selected Topics in Quantum Electronics*, vol. 26, no. 1, p. 7 700 715, 2020. DOI: 10.1109/JSTQE.2019.2931215.
- [102] R. A. Fisher, “The use of multiple measurements in taxonomic problems,” *Annals of Eugenics*, vol. 7, no. 2, pp. 179–188, 1936. DOI: 10.1111/j.1469-1809.1936.tb02137.x.
- [103] I. Fischer, G. H. M. van Tartwijk, A. M. Levine, W. Elsässer, E. Göbel, and D. Lenstra, “Fast pulsing and chaotic itinerancy with a drift in the coherence collapse of semiconductor lasers,” *Physical Review Letters*, vol. 76, no. 2, pp. 220–223, 1996. DOI: 10.1103/PhysRevLett.76.220.
- [104] I. Guyon, J. Li, T. Mader, P. A. Pletscher, G. Schneider, and M. Uhr, “Competitive baseline methods set new standards for the nips 2003 feature selection benchmark,” *Pattern Recognition Letters*, vol. 28, no. 12, pp. 1438–1444, 2007. DOI: 10.1016/j.patrec.2007.02.014.
- [105] E. Benhamou and V. Melot, “Seven proofs of the pearson chi-squared independence test and its graphical interpretation,” 2018. [Online]. Available: <https://arxiv.org/abs/1808.09171>.
- [106] M. C. Mackey and L. Glass, “Oscillation and chaos in physiological control systems,” *Science*, vol. 197, no. 4300, pp. 287–289, 1977. DOI: 10.1126/science.267326.
- [107] L. Jaurigue, E. Robertson, J. Wolters, and K. Lüdge, “Reservoir computing with delayed input for fast and easy optimisation,” *Entropy (Basel, Switzerland)*, vol. 23, no. 12, p. 1560, 2021. DOI: 10.3390/e23121560.
- [108] D. J. Gauthier, E. Bollt, A. Griffith, and W. A. S. Barbosa, “Next generation reservoir computing,” *Nature communications*, vol. 12, no. 1, pp. 5564–5564, 2021. DOI: 10.1038/s41467-021-25801-2.
- [109] T. Hülser, F. Köster, K. Lüdge, and L. Jaurigue, “Deriving task specific performance from the information processing capacity of a reservoir computer,” *Nanophotonics (Berlin, Germany)*, vol. 12, no. 5, pp. 937–947, 2023. DOI: 10.1515/nanoph-2022-0415.

References

- [110] D. Bohm, *Quantum theory* (Prentice-Hall physics series). New York: Prentice-Hall, 1951.
- [111] M. Hejda et al., “Artificial optoelectronic spiking neuron based on a resonant tunnelling diode coupled to a vertical cavity surface emitting laser,” *Nanophotonics*, vol. 12, no. 5, pp. 857–867, 2023. DOI: 10.1515/nanoph-2022-0362.
- [112] B. Romeira, L. M. Pessoa, H. M. Salgado, C. N. Ironside, and J. M. Figueiredo, “Photo-detectors integrated with resonant tunneling diodes,” *Sensors (Basel, Switzerland)*, vol. 13, no. 7, pp. 9464–9482, 2013. DOI: 10.3390/s130709464.
- [113] W. Zhang, A. Al-Khalidi, J. Figueiredo, Q. R. A. Al-Taai, E. Wasige, and R. H. Hadfield, “Analysis of excitability in resonant tunneling diode-photodetectors,” *Nanomaterials*, vol. 11, no. 6, p. 1590, 2021. DOI: 10.3390/nano11061590.
- [114] Q. R. A. Al-Taai et al., “Optically-triggered deterministic spiking regimes in nanostructure resonant tunnelling diode-photodetectors,” *Neuromorphic computing and engineering*, vol. 3, no. 3, p. 34012, 2023. DOI: 10.1088/2634-4386/acf609.
- [115] P. Podziemski and J. Żebrowski, “Liénard-type models for the simulation of the action potential of cardiac nodal cells,” *Physica D: Nonlinear Phenomena*, vol. 261, pp. 52–61, 2013. DOI: 10.1016/j.physd.2013.06.007.
- [116] R. de Figueiredo, “Existence and uniqueness results for lienard’s equations,” *IEEE Transactions on Circuit Theory*, vol. 17, no. 3, pp. 313–321, 1970. DOI: 10.1109/TCT.1970.1083127.
- [117] B. van der Pol Jun., “On “relaxation-oscillations”,” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 978–992, 1926. DOI: 10.1080/14786442608564127.
- [118] J. Schulman, H. De Los Santos, and D. Chow, “Physics-based rtd current-voltage equation,” *IEEE Electron Device Letters*, vol. 17, no. 5, pp. 220–222, 1996. DOI: 10.1109/55.491835.

References

- [119] B. Romeira, J. Javaloyes, C. N. Ironside, J. M. L. Figueiredo, S. Balle, and O. Piro, “Excitability and optical pulse generation in semiconductor lasers driven by resonant tunneling diode photo-detectors,” *Opt. Express*, vol. 21, no. 18, pp. 20 931–20 940, 2013. DOI: 10.1364/OE.21.020931.
- [120] M. Hejda et al., “Resonant tunneling diode nano-optoelectronic excitable nodes for neuromorphic spike-based information processing,” *Phys. Rev. Appl.*, vol. 17, p. 024 072, 2 2022. DOI: 10.1103/PhysRevApplied.17.024072.
- [121] Q. Liu, A. Seabaugh, P. Chahal, and F. Morris, “Unified ac model for the resonant tunneling diode,” *IEEE transactions on electron devices*, vol. 51, no. 5, pp. 653–657, 2004. DOI: 10.1109/TED.2004.825795.
- [122] G. Donati et al., “Spiking flip-flop memory in resonant tunneling diode neurons,” *Phys. Rev. Lett.*, vol. 133, p. 267 301, 26 2024. DOI: 10.1103/PhysRevLett.133.267301.
- [123] J. Mork and G. L. Lippi, “Rate equation description of quantum noise in nanolasers with few emitters,” *Applied Physics Letters*, vol. 112, no. 14, p. 141 103, 2018. DOI: 10.1063/1.5022958.
- [124] P. Virtanen et al., “Scipy 1.0: Fundamental algorithms for scientific computing in python,” *Nature Methods*, vol. 17, pp. 261–272, 2020. DOI: 10.1038/s41592-019-0686-2.
- [125] E. Süli, *An introduction to numerical analysis*. Cambridge University Press, 2003.
- [126] W. Zhang et al., “Photonic-electronic spiking neuron with multi-modal and multi-wavelength excitatory and inhibitory operation for high-speed neuromorphic sensing and computing,” *Neuromorphic computing and engineering*, vol. 4, no. 4, p. 44 006, 2024. DOI: 10.1088/2634-4386/ad8df8.
- [127] M. Carandini, F. Mechler, C. S. Leonard, and J. A. Movshon, “Spike train encoding by regular-spiking cells of the visual cortex,” *Journal of Neurophysiology*, vol. 76, no. 5, pp. 3425–3441, 1996, PMID: 8930283. DOI: 10.1152/jn.1996.76.5.3425.

References

- [128] D. Auge, J. Hille, E. Mueller, and A. Knoll, “A survey of encoding techniques for signal processing in spiking neural networks,” *Neural processing letters*, vol. 53, no. 6, pp. 4693–4710, 2021. DOI: 10.1007/s11063-021-10562-2.
- [129] M. Borghi, S. Biasi, and L. Pavesi, “Reservoir computing based on a silicon microring and time multiplexing for binary and analog operations,” *Scientific Reports*, vol. 11, no. 1, p. 15 642, 2021. DOI: 10.1038/s41598-021-94952-5.
- [130] J. R. Rausell-Campo, A. Hurtado, D. Pérez-López, and J. Capmany Francoy, “Programmable photonic extreme learning machines,” *Laser & photonics reviews*, vol. 19, no. 9, 2025. DOI: 10.1002/lpor.202400870.
- [131] S. Biasi, R. Franchi, L. Cerini, and L. Pavesi, “An array of microresonators as a photonic extreme learning machine,” *APL photonics*, vol. 8, no. 9, pp. 096 105–096105-12, 2023. DOI: 10.1063/5.0156189.
- [132] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997. DOI: 10.1162/neco.1997.9.8.1735.
- [133] J. M. Martins, S. V. Gurevich, and J. Javaloyes, “Excitability and memory in a time-delayed optoelectronic neuron,” *Phys. Rev. Appl.*, vol. 22, p. 024 050, 2 2024. DOI: 10.1103/PhysRevApplied.22.024050.
- [134] B. Romeira, R. Avó, J. M. L. Figueiredo, S. Barland, and J. Javaloyes, “Regenerative memory in time-delayed neuromorphic photonic resonators,” *Scientific Reports*, vol. 6, p. 19 510, 1 2016. DOI: 10.1038/srep19510.
- [135] S. Terrien et al., “Equalization of pulse timings in an excitable microlaser system with delay,” *Physical review research*, vol. 2, no. 2, p. 023 012, 2020. DOI: 10.1103/PhysRevResearch.2.023012.
- [136] S. Terrien, V. A. Pammi, B. Krauskopf, N. G. R. Broderick, and S. Barbay, “Pulse-timing symmetry breaking in an excitable optical system with delay,” *Physical review. E*, vol. 103, no. 1, pp. 012 210–012 210, 2021. DOI: 10.1103/PhysRevE.103.012210.

References

- [137] S. Boyd and L. Chua, “Fading memory and the problem of approximating non-linear operators with volterra series,” *IEEE transactions on circuits and systems*, vol. 32, no. 11, pp. 1150–1161, 1985. DOI: 10.1109/TCS.1985.1085649.
- [138] J. Dambre, D. Verstraeten, B. Schrauwen, and S. Massar, “Information processing capacity of dynamical systems,” *Scientific Reports*, vol. 2, p. 514, 1 2012. DOI: 10.1038/srep00514.
- [139] J. Robertson, T. Deng, J. Javaloyes, and A. Hurtado, “Controlled inhibition of spiking dynamics in VCSELs for neuromorphic photonics: theory and experiments,” *Opt. Lett.*, vol. 42, no. 8, pp. 1560–1563, 2017. DOI: 10.1364/OL.42.001560.
- [140] M. Hejda, J. Robertson, J. Bueno, J. A. Alanis, and A. Hurtado, “Neuromorphic encoding of image pixel data into rate-coded optical spike trains with a photonic VCSEL-neuron,” *APL Photonics*, vol. 6, no. 6, p. 60 802, 2021. DOI: 10.1063/5.0048674.
- [141] J. Robertson, Y. Zhang, M. Hejda, J. Bueno, S. Xiang, and A. Hurtado, “Image edge detection with a photonic spiking VCSEL-neuron,” *Optics Express*, vol. 28, no. 25, pp. 37 526–37 537, 2020. DOI: 10.1364/OE.408747.
- [142] A. Grabulosa, J. Moughames, X. Porte, M. Kadic, and D. Brunner, “Additive 3d photonic integration that is cmos compatible,” eng, *Nanotechnology*, vol. 34, no. 32, p. 322 002, 2023.