

Neuromorphic nanophotonic systems for artificial intelligence

PhD Thesis

Matěj Hejda

Primary supervisor: Antonio Hurtado
Institute of Photonics, Dept. of Physics
University of Strathclyde, Glasgow

November 21, 2023

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

Over the last decade, we have witnessed an astonishing pace of development in the field of artificial intelligence (AI), followed by proliferation of AI algorithms into virtually every domain of our society. While modern AI models boast impressive performance, they also require massive amounts of energy and resources for operation. This is further fuelling the research into AI-specific, optimised computing hardware. At the same time, the remarkable energy efficiency of the brain brings an interesting question: *Can we further borrow from the working principles of biological intelligence to realise a more efficient artificial intelligence?* This can be considered as the main research question in the field of neuromorphic engineering.

Thanks to the developments in AI and recent advancements in the field of photonics and photonic integration, research into light-powered implementations of neuromorphic hardware has recently experienced a significant uptick of interest. In such hardware, the aim is to seize some of the highly desirable properties of photonics not just for communication, but also to perform computation. Neurons in the brain frequently process information (compute) and communicate using *action potentials*, which are brief voltage spikes that encode information in the temporal domain. Similar dynamical behaviour can be elicited in some photonic devices, at speeds multiple orders of magnitude higher. Such devices with the capability of neuron-like spiking are of significant research interest for the field of neuromorphic photonics. Two distinct types of such excitable, spiking systems operating with optical signals are studied and investigated in this thesis.

First, a vertical cavity surface emitting laser (VCSEL) can be operated under a specific set of conditions to realise a high-speed, all-optical excitable photonic neuron that operates at standard telecom wavelengths. The photonic VCSEL-neuron was

dynamically characterised and various information encoding mechanisms were studied in this device. In particular, a spiking rate-coding regime of operation was experimentally demonstrated, and its viability for performing spiking domain conversion of digital images was explored. Furthermore, for the first time, a joint architecture utilising a VCSEL-neuron coupled to a photonic integrated circuit (PIC) silicon microring weight bank was experimentally demonstrated in two different functional layouts.

Second, an optoelectronic (O/E/O) circuit based upon a resonant tunnelling diode (RTD) was introduced. Two different types of RTD devices were studied experimentally: a higher output power, μ -scale RTD that was RF coupled to an active photodetector and a VCSEL (this layout is referred to as a PRL node); and a simplified, photosensitive RTD with nanoscale injector that was RF coupled to a VCSEL (referred to as a nano-pRL node). Hallmark excitable behaviours were studied in both devices, including excitability thresholding and refractory periods. Furthermore, a more exotic resonate-and-fire dynamical behaviour was also reported in the nano-pRL device. Finally, a modular numerical model of the RTD was introduced, and various information processing methods were demonstrated using both a single RTD spiking node, as well as a perceptron-type spiking neural network with physical models of optoelectronic RTD nodes serving as artificial spiking neurons.

Contents

Abstract	ii
List of Figures	vii
List of Tables	xxiii
List of Abbreviations	xxv
Acknowledgements	xxviii
1 Introduction	1
1.1 Motivation	1
1.2 Overview of key concepts	2
1.2.1 Artificial neural networks (ANNs)	2
1.2.2 Spiking neural networks (SNNs)	4
1.2.3 Nonlinear dynamics and excitability	6
1.2.4 Dynamics of biological neurons	8
1.2.5 Semiconductor lasers	10
1.3 Description of the components and equipment used in the experiments .	11
1.3.1 Laser driving components	12
1.3.2 Fibre-optic components	12
1.3.3 Electronics equipment used in the experimental work of this thesis	14
1.3.4 Radio frequency (RF) probing equipment	15
1.4 Thesis outline	16

2	Neuromorphic photonics and optical neural networks: an overview	19
2.1	Introduction	19
2.1.1	Motivation for alternative computing approaches	21
2.2	Neuromorphic engineering	26
2.2.1	Definition	27
2.2.2	Characteristics of neuromorphic systems	28
2.2.3	Neuromorphic electronic platforms	29
2.3	From electrons to photons: Neuromorphic photonics	32
2.3.1	Advantages of photonics for neuromorphic systems	33
2.3.2	Disadvantages of photonics for neuromorphic systems	34
2.3.3	Taxonomy of light-powered neuromorphic approaches	35
2.4	Photonic neural networks: Integrated approaches	36
2.4.1	Mach-Zehnder interferometers (MZIs)	36
2.4.2	Semiconductor optical amplifiers (SOAs)	38
2.4.3	Microring resonators (MRRs)	39
2.4.4	Electro-absorption modulators (EAMs)	41
2.4.5	Phase-change materials (PCMs)	42
2.4.6	Josephson junction-based optoelectronics	44
2.5	Photonic neural networks: Spatial approaches	45
2.6	Photonic neural networks: Laser-based approaches	45
3	Photonic neurons based on VCSELs: characterisation and applications	47
3.1	Fundamentals of VCSELs	48
3.1.1	VCSEL characteristics	48
3.1.2	Injection locking of VCSELs	50
3.2	Spiking in injection-locked VCSEL neurons	51
3.2.1	State-of-the-art in VCSEL-neurons	53
3.2.2	VCSELs used in this thesis	55
3.2.3	Experimental methods	55
3.3	Characterisation of excitable spiking dynamics in the VCSEL-neuron	57
3.3.1	Spike shape characterisation	58

3.3.2	Refractory period characterisation	59
3.3.3	Optical spiking trigger latency evaluation	61
3.3.4	Demonstration of leaky integrate-and-fire functionality	64
3.4	Experimental information encoding	66
3.4.1	Optical digital-to-spike (DtS) format conversion	66
3.4.2	Continuous signal encoding	68
3.4.3	Neuromimetic rate-coding of digital images	70
3.5	Interfacing VCSEL-neurons with SiPh weight bank chips	81
3.5.1	VCSEL-neuron spike weighting in SiPh MRR	83
3.5.2	Rate-coding in SiPh MRR-VCSEL-neuron system	88
4	Artificial spiking neurons based upon resonant tunnelling diodes (RTDs)	93
4.1	Overview, physics and band structure of RTDs	94
4.2	Modelling (opto)electronic RTD circuits	97
4.2.1	van der Pol & Liénard oscillators	97
4.2.2	Nonlinear I - V characteristic of the RTDs	99
4.2.3	Influence of RTD circuit parameters	100
4.2.4	Excitability in RTDs	102
4.2.5	Extension to an optoelectronic model	105
4.3	Numerical analysis of a monolithic O/E RTD (PRL-MN)	108
4.3.1	Parameters of the PRL-MN model	109
4.3.2	Spiking and excitable threshold in the PRL-MN model	113
4.3.3	Pre-synaptic spike weighting in the PRL-MN model	115
4.4	Experimental analysis of μ -RTDs (PRL nodes) for spiking neuronal models	116
4.4.1	Device characteristics, experimental setup: PD-RTD-VCSEL (PRL)	117
4.4.2	Characterisation of the spiking responses in the PRL node	121
4.4.3	Pre-synaptic spike weighting in the PRL	126
4.5	Experiments with nano-pRTDs (n-pRL nodes)	127
4.5.1	Device, experimental setup	127
4.5.2	Dynamical characterisation: pRTD	130

4.5.3	Dynamical characterisation: coupled spiking nano-pRTD-LD . . .	134
4.5.4	Resonate-and-fire behaviour	137
5	Applications and networks of RTD-powered neuromorphic nodes	141
5.1	Information processing in a numerical PRL-MN node	142
5.1.1	Time-domain multiplexed pattern recognition	142
5.1.2	Sequential convolution-like operation for feature detection	145
5.2	Spiking optoelectronic neural network with RTD-MN nodes	148
5.2.1	Hybrid supervised-STDP learning in the RTD-SpikeSLP	150
5.2.2	Inference in the RTD-SpikeSLP model	154
5.2.3	Discussion of the RTD-SpikeSLP model	155
5.3	Experimental analysis of information processing in a PRL node	156
5.3.1	Coincidence detection in a single PRL node	156
5.3.2	Exclusive OR (XOR) task in a single PRL node	157
5.4	Dynamically-weighted optical interlink of RTD spiking nodes	159
6	Conclusions and Outlook	165
6.1	Thesis summary	165
6.1.1	All-optical artificial spiking VCSEL-neurons	165
6.1.2	Optoelectronic artificial RTD-neurons	167
6.2	Future outlooks	169
6.2.1	Device level	170
6.2.2	Functional interlinking	171
6.2.3	Operational protocols	172
6.2.4	System-level approach	173
	List of scientific contributions	177
	List of references	183

List of Figures

1.1	Excerpt of "The Neural Network Zoo", demonstrating various architectures of artificial neural networks. Reproduced and modified from [3].	3
1.2	Simplified schematic of a general spiking neural network.	4
1.3	Comparison of spike firing mechanisms in integrate-and-fire and resonate-and-fire neuronal dynamical models. Adapted from [40].	9
1.4	Simplified schematic of a semiconductor laser (laser diode) with a Fabry-Perot type cavity. Reproduced and adapted from [45].	10
2.1	Illustration of growth in compute demands for AI models. During the 'AI golden era' of the past decade, rate of growth in compute requirements of new models has significantly outpace the historical norm. Reproduced from [65].	22
2.2	Selected key performance metrics of current digital computers, demonstrating stalling trends in performance. Available from [77].	25
2.3	Comparison between top-down and bottom up approaches to designing AI-specific computing architectures.	28
2.4	Keyword (research motivation) trends in publications over time. Reproduced from [94].	30
2.5	Illustration of the optical interference unit, which realises all-optically both the functions of matrix multiplication (highlighted in red) and attenuation (highlighted in blue). Reproduced from [141].	37

2.6	Experimental setup of SOA-based optical neural network. The Transmitter block shows how the four signals are modulated and multiplexed prior to entering the PIC. PCs: polarisation controllers; DMUX: Demultiplexer; Pre-SOA: Semiconductor optical preamplifier; VS: Input vector selector gate; AWG: Array waveguide grating; EDFA: Erbium doped fibre amplifier; PD: Photodetector; OSA: Optical spectrum analyzer; DPO: Oscilloscope (OSC). Tunable time delay indicated with dashed lines. Reproduced from [169].	39
2.7	Silicon photonic non-spiking neuron implementation using a microring resonator weight bank. Reproduced from [176].	40
2.8	An electro-optic neuron layout. Input is provided by (a) an array of rings or (b) with an interferometer network. The neuron sums the optical signal with a photodiode converting the signal to a voltage (c), optionally amplified by a TIA (d), drives an electro-optic modulator (e) modulating a CW laser (f), which produces a nonlinear transfer function at the output (g) of an photonic neural network (NN). Reproduced from [184].	42
2.9	Scheme of the multiplexed all-optical PCM-enabled neural network (MAC unit). The input vectors are using a comb source. The entries of different input vectors are grouped together again employing wavelength multiplexing and sent to the on-chip MAC unit that performs the calculations. After combining the correct wavelengths with optical wavelength division demultiplexers, the multiplication results are obtained from the photodetectors followed by digital signal processing. Reproduced from [189].	43

2.10	(A) A deep diffractive processing unit is based on multiple transmissive (or reflective) layers, where each point on a given layer acts as a neuron, with a complex-valued transmission (or reflection) coefficient. The transmission or reflection coefficients of each layer can be trained by using deep learning to perform a function between the input and output planes of the network. After learning, the layers are 3D-printed and perform a specific task at the speed of light. (B) visualisation of a classifier for handwritten digits. Reproduced from [200].	46
3.1	A schematic of a typical VCSEL device structure, consisting of an active semiconductor medium between two DBRs. Reproduced from [210]. . .	50
3.2	Diagram of the experimental setup used to investigate the VCSEL-neuron. TL – tunable laser; ISO – optical isolator; VOA – variable optical attenuator; PC – polarisation controller; 50/50 – optical splitter; PM – optical power meter; CIRC – optical circulator; OSA – optical spectrum analyzer; PD – photodetector; RT OSC – real-time oscilloscope; AWG – arbitrary waveform generator. Reproduced from [254].	56
3.3	(Left) Optical spectrum of a free-running VCSEL biased with current of $I_{VCSEL} = 5$ mA. The spectrum exhibits two polarisation modes referred to as λ_y (parallel) and λ_x (orthogonal). Reproduced from [255]. (Right) LI curve of the used 1310nm VCSEL.	58
3.4	Excitable spiking in the VCSEL neuron under orthogonally-polarised optical injection. (a) Optical input signal with the encoded modulated perturbations to elicit spike events. (b) Parallel and (c) orthogonal polarisation-filtered signal collected from the output of the modulated VCSEL-neuron. (d) Composite signal of both polarisations, as typically recorded from the VCSEL-neuron output. The estimated FWHM of the output neuron-like excitable spikes is reported.	59

3.5	Refractory period measurement in the VCSEL-neuron. Upper (red) time traces show pairs of input perturbations with increasing temporal separation. The second perturbation arrives (a) within the absolute refractory period (red shading), (b) within the relative refractory period (orange shading), (c) after the refractory period, therefore eliciting a second spike. (d) Probability of firing a second optical spike versus temporal separation between input perturbations over $n = 70$ consecutive measurement cycles. Reproduced from [255].	60
3.6	Demonstration of optical spiking latency in the VCSEL-neuron. (a) Variation in spike onset latency and trigger irregularity (jitter) as function of trigger amplitude. Overlay of $n = 5$ measurements. (b) Temporal maps of $n = 150$ measurement cycles, showing responses to 20% (left) and 90% (right) strength perturbations. (c) Mean spike onset delay time as a function of perturbation strength. Total of $n = 150$ repetitions, the last spike is used as timing reference. Reproduced from [255].	62
3.7	Demonstration of integrate-and-fire spiking responses in the VCSEL-neuron. The VCSEL-neuron is operated (a) further, (b) closer to the excitability threshold by adjusting injection frequency detuning. Red traces represent an example of the modulated injected optical signal (same demonstrative trace shown in both figures), while blue represent responses from the VCSEL-neuron.; Temporal maps constructed over $n = 750$ repetitions that plot the continuous time traces with the spiking and quiescent responses from the VCSEL-neuron represented in red and blue colours, respectively.	65
3.8	a) Simplified schematic demonstrating the operation of a photonic VCSEL-neuron. b) Working principle of the two distinct information encoding regimes in the VCSEL-neuron: precise temporal coding and rate coding.	66

3.9	Digital-to-spike information encoding in the VCSEL-neuron for return-to-zero (RZ) input code. (a) (top, in red) Input sequence of 24 bits representing in binary the word "IoP", an acronym of the Institute of Photonics; (bottom, in blue) superposition of five time series recorded at the output of the VCSEL-neuron. (b) Temporal map of $n = 150$ consecutive repetitions of time traces recorded from VCSEL-neuron. (c) Standard deviations of timing for each spike across n repetitions. Reproduced from [255].	67
3.10	Continuous spiking rate-coding in a VCSEL-neuron, using (a,b) a sawtooth wave modulation, and (c,d) sinusoidal modulation. (a,c) Single traces for each modulation case, with modulation shown in red and VCSEL-neuron response in blue. (b,d) Temporal maps over $n = 150$ measurements, with histograms (1 ns bins).	69
3.11	Digital image channel data pre-processing prior to rate-encoding in the VCSEL-neuron. Reproduced from [254].	72
3.12	Emission spectrum of the VCSEL at $I = 5$ mA, with polarisation-selective filtering to demonstrate the two polarisations of the main transversal model. Reproduced from [254].	73
3.13	Spike-rate encoding of an 8×8 GS image with a VCSEL-neuron. (a) Output power levels from the MZM for each of the 64 power drops when modulating with waveform from Fig. 3.11(c).; (b) Temporal maps created from the experimental data, showing the response of VCSEL-neuron during each modulation power drop of $T_p = 60$ ns (T_0 interleave period is omitted). (c) Mean spiking rates calculated for each cycle. (d) Spiking rate produced by the VCSEL-neuron as a function of input perturbation strength (drop amplitude). Reproduced from [254].	75

- 3.14 Spike-rate encoding of an 8×8 , two-channel RB image with a VCSEL-neuron. Results are presented for the individual RB-R (a-d) and RB-B (e-h) colour channels: (a, e) For each of the 64 pixels (cycle) in the RB-R (a) and RB-B channels (e) in the RB diagonal source image, the pixel intensity is encoded in the magnitude of the input perturbation (power drop) added to the optical signal injected into the VCSEL-neuron, with the estimated spiking threshold (\mathcal{T}_{est}) highlighted as a red line; (b, f) Temporal maps for the responses of the RB-R (b) and the RB-B channels (f); (c, g) Calculated average spiking rates for each cycle (pixel) in the RB-R (c) and RB-B channels (g); (d, h) Spiking rate produced by the VCSEL-neuron as a function of input perturbation strength (drop amplitude) in the RB-R (d) and RB-B channels (h). Reproduced from [254]. 77
- 3.15 Source images and reconstructions based on the spike-encoded representation for GS and RB-R, RB-B channels. Source colour channels are shown in upper row, and channels directly reconstructed from the spike rates are shown in bottom row. Reproduced from [254] 78
- 3.16 Mean obtained optical spike firing rates from the VCSEL-neuron as a function of the input pixel intensity values for each colour channel in the RGB image: a) red (R); b) green (G); c) blue (B). Each colour dot in the plots corresponds to an individually processed pixel. Mean values per each individual pixel value are shown as black dots, with error bars corresponding to single standard deviation. The sparsity observed in the red channel comes from the red colour distribution in the source image. Reproduced from [254]. 80

3.17	Processing of a complex, pixel-art 32×32 RGB image. (<i>upper row</i>) Individual R/G/B channels of the image, with source image shown at left. (<i>lower row</i>) Rate-encoded and reconstructed R/G/B channels produced by the VCSEL-neuron, with composite reconstruction image shown at left. The contrast of the individual reconstructed channels and directly reconstructed composite image was adjusted to better match to the source images and allow for clearer comparison. Reproduced from [254].	81
3.18	Characterisation of the MRR within the integrated SiPh weight bank. Upper three plots (a,b,c) correspond to characteristics recorded on the THRU port of the MRR, while (d,e,f) correspond to readout from the DROP port readouts of the MRR. (b,e) Wavelength of MRR resonance as a function of V_{MR} . (c,f) Recorded optical power at a fixed wavelength (dashed line) and maxima (minima) position as function of applied MRR bias V_{MR}	83
3.19	VCSEL-to-MRR experimental setup. TL - tuneable laser; ISO - isolator; VOA - variable optical attenuator; PC(1,2,3) - polarisation controllers; MZM - Mach-Zehnder modulators; PM - power meter; CIRC - circulator; VCSEL - vertical cavity surface emitting laser; EDFA - erbium-doped fibre amplifier; BPF - bandpass filter; RT OSC - real-time oscilloscope; OSA - optical spectrum analyzer; GC - grating coupler; PD - photodetector. Ratios for the optical couplers/splitters are indicated in the image. . .	84
3.20	Optical spectra of the VCSEL when operated in (green) a free-running, CW lasing regime; (red) under the external injection of an optical signal into the device's subsidiary mode, resulting in injection locking with associated polarisation switching. This injection locked trace was acquired while operating the device in a constantly spiking regime, hence widening the shape of the spectrum.	85

3.21	Sequences of deterministically triggered excitable spikes in the VCSEL-neuron after passing through the weighting MRR for three different values of V_{MR} . Traces were recorded on both outputs of the micro-ring: the THRU port (left column, in red) and the DROP port (right column, in green).	86
3.22	Figure of merit: spike amplitude modulation via MRR voltage tuning. (DC floor is removed).	87
3.23	The figure of merit for spike weighting: spike amplitude modulation performed via MRR resonance tuning, as recorded from the THRU (in red) and DROP (in green) ports of the MRR. This tuning is performed by varying a bias V_{MRR} on a thermo-optic heater of the MRR. (The DC floor is removed during data analysis).	88
3.24	MRR \rightarrow VCSEL rate-coding experimental setup. TL - tuneable laser; VOA - variable optical attenuator; PC1 - polarisation controllers; PM - power meter; CIRC - circulator; VCSEL - vertical cavity surface emitting laser; EDFA - erbium-doped fibre amplifier; BPF - bandpass filter; RT OSC - real-time oscilloscope; OSA - optical spectrum analyzer; GC - grating coupler; PD - photodetector. Ratios for the power splitters/couplers are indicated in the figure.	89
3.25	Demonstration of rate coding in a spiking VCSEL-neuron subject to injection of optical signals from a MRR. (a) Injection signal power as a function of V_{MR} . (b) Spike firing rate in the VCSEL-neuron as a function of V_{MR} . (c) Mean spike amplitudes within the optical spike train. (d-f) Examples of continuous spiking traces produced by the VCSEL-neuron locked to the signal from the MRR. As the voltage applied to the MRR heater is gradually increased from (a) 1.5 V through (b) 1.95 V to (c) 2.25 V, the spike firing rate monotonically increases.	90

4.1	Simplified scheme of the resonant tunnelling diode (RTD) bandstructure, showing the double barrier quantum well (DBQW) sandwiched between two narrower-bandgap materials. a) No external bias is present. b) By applying an external voltage, the additional potential shifts the bands and aligns the resonant well state with the conduction band, allowing for resonant tunnelling through the DBQW, increasing the measured current. c) Once the applied voltage shifts the bands even further, the well state gets misaligned with respect to the conduction band, tunnelling probability drops to \approx zero and only the diode current is present, resulting in a drop of measured current in the I - V characteristic.	95
4.2	(left) Simplest RTD electronic circuit scheme. (right) Example of an RTD circuit with additional branch of shunting elements in orange. Equivalent elements of the RTD are highlighted in green. $f(V)$ denotes the nonlinear I - V	98
4.3	Comparison of an idealised RTD I - V (third-order polynomial, in black) with a more realistic I - V (in red). Reproduced and adapted from [291].	99
4.4	Dynamical evolution of an RTD circuit biased in the valley region of its I - V curve ($V_0 = 730$ mV) subject to a single voltage perturbation (a square pulse of 100 ps length and -125 mV amplitude) for variations of capacitance (along x -axis) and inductance (along y -axis).	103
4.5	Current-voltage (I - V) characteristic of the modelled RTD device with the NDR highlighted (in orange).	110
4.6	Scanning I-V curves of an RTD. The starting values correspond to a nanoscale RTD device, and parameter value scans are performed between 75% and 150% of starting values. In the case shown in (f), no significant variation is observed for N2 under the current set of parameters.	111
4.7	Thresholding demonstration in the PRL-MN numerical model, with traces (top to bottom) demonstrating input perturbations (red for optical, green for electrical), voltage, current, photon count and carrier number respectively.	114

4.8	Phase space plots for gradually increasing optical perturbations in the valley operation point of the RTD. The asterisk in the right-most figure represents the moment when a full excitable spike is elicited in response to a sufficiently strong perturbation.	114
4.9	Demonstration of spike weighting in the numerical RTD-MN model. (a-j) Time traces of PRL-MN optical output, showing single optical spike for variations in J . (k) Mean spike amplitudes (signal floor subtracted) obtained over $n = 20$ repetitions. (l) Signal-to-noise (SNR) for obtained spikes.	116
4.10	Current-voltage (I - V) characteristic of the used micro-RTD device (forward biased) with the NDR region highlighted (in orange).	117
4.11	A microscope image of the wafer with μ -RTD devices used in the experiments. These were fabricated by E. Malysheva (V. Dolores-Calzadilla's group) at the Eindhoven University of Technology (TUe).	119
4.12	Experimental setup for the characterisation of the opto-electronic dynamical spiking regimes in the module PRL architecture.	119
4.13	Photograph of the probe station used to drive RTD circuits. Here, a two-terminal device is probed using RF probes attached to micro-positioning systems.	120
4.14	Experimental results demonstrating deterministic excitable spike firing in the O/E/O PRL node. (a) RF signal with encoded perturbation generated by the AWG which is then converted to an optical input pulse (b) using an optical modulator. (c) Electrical spike event fired by the RTD in response to the input optical perturbation. (d) Final output optical spike from the PRL node measured after the VCSEL in its structure.	122

4.15	Experimental demonstration of spike firing threshold in the PRL optoelectronic spiking neuron. (a) Source RF waveform (from the AWG). Dashed lines approximate the spike firing threshold in relation to V_0 . (b)-(d) PRL node output, as recorded on an amplified photodetector. For the lowest bias voltage value (b) V_0 860 mV, only the strongest perturbation triggers a response (an optical spike). Response of the system for the same input signal but with the RTD biased now with V_0 870 mV (c) and with V_0 880 mV (d). Reproduced from [267].	123
4.16	Experimental demonstration of refractoriness in the PRL spiking node: (a) Source modulation (input signal into the PRL) consisting of pulse pairs with gradually increasing temporal separation (in 10 ns increments). (b) Temporal map plotting the measured time-traces of 20 subsequently recorded measurement cycles of 10 μ s length. (c) Example of one measured time-trace from the PRL node in response to the input signal in (a). Reproduced from [267].	125
4.17	Weighting of the PRL output spike amplitude via control of the VCSEL bias current (below its lasing threshold). (a)-(h) Optical output spike amplitude when the bias current is changed from 1.1 mA to 1.8 mA with a step size of 100 μ A. In addition, individual time traces of the PRL with VCSEL biased at (i) 1.1 mA and (j) 1.8 mA. The trend is confirmed by (k) mean peak speak amplitudes over 5 repetitions as a function of the VCSEL bias, including a linear fit.	126
4.18	(a) Scheme of the photodetecting RTD with nano-injector (nano-pRTD). (b,c) SEM images of the upper, nanoscale electrode of the nano-pRTD. Reproduced from [319].	129
4.19	Photograph of the die containing the pRTD devices. This sample was fabricated Dr Qusay R. A. Al-Taai and team of Prof. Edward Wasige at the University of Glasgow.	130
4.20	Simplified scheme of the experimental setup used to operate the nano-pRTD spiking neuron.	131

4.21	<i>I-V</i> characteristic of a solitary nano-pRTD device (reverse biased) under darkness (black dashed line) and when subject to 1 mW CW IR laser light illumination (solid red line).	131
4.22	Characterisation of the spiking dynamical behaviour of the nano-pRTD node when biased at the peak ($V_0 = 590$ mV) and valley ($V_0 = 688$ mV) operation points. Repeated deterministic spike triggering elicited in the nano-pRTD when biased at the peak (a) and the valley points (b). Spike thresholding behaviour of the nano-pRTD when biased at peak (c) and valley points (d). Refractory period of the nano-pRTD when biased at the peak (e) and valley points (f). We must note that in order to match the corresponding operational modes at the peak and valley biasing points, the optical input trigger signals (generated from the AWG) are designed with negative/positive polarity.	133
4.23	Simplified scheme of the experimental setup used to operate the nano-pRL optoelectronic spiking neuron. In comparison to Fig. 4.20, there is a laser connected in series at the output.	134
4.24	Evaluation of thresholding in the nano-pRL system. Green time traces show the input RF modulation waveforms, which are used (via the MZM) to modulate CW IR light signal input to the nano-pRL system, while read time traces show the optical output from the VCSEL in the nano-pRL node. Left column shows responses in peak operation points, while right column shows responses in the valley operation points. In both cases, two different RTD bias points (altering the distance of the steady state from the spiking threshold) have been evaluated.	135

4.25	Evaluation of refractoriness in the nano-pRL system. Green time traces show the input RF modulation waveforms, which are used (via the MZM) to modulate CW IR light signal input to the nano-pRL system, while read time traces show the optical output from the VCSEL in the nano-pRL node. Left column shows responses in the peak operation point ($V_0 = 630$ mV), while right column shows responses in the valley operation point ($V_0 = 663$ mV).	136
4.26	Resonate-and-fire effect in the pRTD. The green time trace (top) shows the modulation waveform (used in MZM for optical input into the pRTD) with pairs of sub-threshold perturbations. Purple (middle) trace shows overlay of $n = 50$ pRTD electrical output time traces. Dashed (pink) bottom time trace shows the normalised occurrence rate of pRTD spiking for each pair of input sub-threshold perturbations.	138
5.1	Sequential data processing for the recognition of serialised Tetris-like input patterns, realised using the PRL-MN spiking optoelectronic neuron model. Only for the case where the input pattern P matches the kernel K , the pattern is pre-processed into a sequence of input pulses that triggers and excitable spike in the RTD. (a) Input pulse sequences. (b) PRL-MN voltage response. (c) PRL-MN current response. (d) PRL-MN optical response.	144

5.2	Processing steps and example of the convolution-like feature extraction (edge detection) in the PRL-MN. a) Schematic showing the individual steps of data processing in the PRL-MN model. A sliding window approach is utilised, followed by a MAC operation for each window position, serialisation, sorting and dual-polarity PAM modulation. This is fed into the optoelectronic RTD model, which may respond with optical spikes. b)-d) Shows individual steps for processing (edge feature detection) for an example of a black-and-white cross image. e) Cross image with extracted features. f) Second example, using larger image with Institute of Photonics (IOP) logo. g) IOP logo with features extracted using the same process. Reproduced from [320].	146
5.3	a) Network architecture diagram. The output state of the downstream node is compared to the label, and if there is a mismatch between label and output state, the weights are updated. The desired pattern is highlighted with the target icon. b) visualisation of inference in the 5-to-1 feed-forward network model. The guiding signals representing pattern labels are visualised as background shading (green for 'True', grey for 'False'). Only a particular spatial pattern ([1 0 1 0 1], green) results in firing of an electric spike by the downstream RTD-PD node (green current trace). The red time trace represents a simple moving average of the LD output optical signal over $t_{MA} = 100$ ps. Reproduced from [320].	149
5.4	Demonstration of the supervised learning process for the spatio-temporal patterns. a) [0 1 0 1 0] and b) [1 0 1 0 1]. Randomly sampled labelled patterns are used for the training via the custom learning rule. The background colour shows the network state (True/False) during each step. Reproduced from [320].	152

5.5	Weight adjustment factor Δw_n as a function of POST-PRE spiking interval $ \Delta T_n $. For false negatives (FN, in green), the weight adjustment is a constant fixed value. For false positives, the weight adjustment magnitude is a function of $ \Delta T_n $, with closer spikes yielding stronger depotentiation. Reproduced from [320].	153
5.6	Confusion matrices for <i>a)</i> for inference of pattern $[0\ 1\ 0\ 1\ 0]$ against all other patterns with two ON bits ($n = 10$ different patterns, 540 total inference steps); <i>b)</i> for inference of pattern $[1\ 0\ 1\ 0\ 1]$ against all other patterns with three ON bits ($n = 10$ different patterns, 540 total inference steps).	154
5.7	Two-input coincidence detection (logical AND) realised in the PRL node. In both left and right figures, the first two rows show the 2 ns square pulses (input stimuli) being gradually mutually shifted in time, in 4 different shifts of 0 ns to 3 ns. For perfect (leftmost) and for partial (2nd from the left) overlap, the total energy at a given moment surpasses the firing threshold, eliciting a spike. However, as the pulses drift further apart (c,d), the input perturbation energy to the system at any given time is not sufficient for spiking. Reproduced from [267].	156
5.8	The pre-processing layer for the XOR task using the PRL, realised with MZMs. (a,d) The bipolar trigger pulses are processed through a layer of two (b,e) MZMs with different settings, providing mutually inverted input perturbations (c,f) on two optical input branches (denoted as A,B). Reproduced from [267].	157
5.9	The exclusive OR (XOR) task demonstrated experimentally in the PRL. The input pulses (see Fig. 5.8) can individually elicit spiking responses from the PRL node (g,h), but when arriving simultaneously, they cancel out (i) and no spike is fired (j). This demonstrates the XOR task operation. Reproduced from [267].	158

5.10	Scheme of the experimental setup and layout of the RTD-laser to PD-RTD interlink (or RL-PR in short) system with dynamical (MZM-enabled) weighting. Signals from one arbitrary waveform generator (AWG) channel provide electrical perturbations that activate electrical spiking in the master RTD. These spikes are used as modulation for a VCSEL, providing optical spiking regimes. These optical spikes are dynamically weighted at the MZM using second AWG signal. Finally, weighted spikes are directly injected into the receiver pRTD via a lens-ended fibre. . . .	160
5.11	Demonstration of dynamic pulse weighting in a master-receiver layout of two RTD optoelectronic nodes (PRL \rightarrow pRTD arrangement). (a) The electrical trigger pulses (stimuli) used to elicit excitable spiking in the master μ -RTD. (b) Optical spiking signals from the master RTD (biased in peak region) coupled to a VCSEL (for E/O conversion), in response to the stimuli. (c) Dynamical weighting waveform applied to the synaptic MZM element on the interlink. (d) Dynamically weighted optical spikes from the master RTD (RTD-VCSEL). (e) Responses of the receiver (p-RTD) to the weighted optical spikes, injected directly into the p-RTD via a lens-ended fibre. (f) Histogram of spike elicitation probabilities as a function of spike weighting factor (spike amplitude). . .	162
6.1	Overview of thesis results.	166

List of Tables

3.1	Advantages of VCSELs compared to EELs (left) and communication-type LEDs (right), Reproduced from [204].	49
4.1	Modelled RTD I - V curve parameters	110
4.2	Modelled RTD circuit parameters	112
4.3	Modelled RTD-LD circuit: laser parameters	112
4.4	Epi-layer stack for the μ -RTD device. This device was designed and fabricated by team of our project collaborators at the Eindhoven University of Technology (E. Malysheva, V. Calzadilla).	118
4.5	Epi-layer stack of the nano-pRTD device. This was designed and fabricated by our project collaborators at the University of Glasgow (Dr Qusay Al-Taai, Prof Edward Wasige).	128

List of Abbreviations

Symbols

I-V voltage-current characteristic.

◆ Non-spiking (CW) optical computing hardware.

◻ Spiking optical computing hardware.

A

ADC analog-to-digital converter.

AI Artificial Intelligence.

ANN artificial neural network.

ASIC application-specific integrated circuit.

AWG arbitrary waveform generator.

B

BPF band-pass filter.

C

CAGR compound annual growth rate.

CIRC circulator.

CMOS complementary metal-oxide-semiconductor.

CNN convolutional neural network.

CPU central processing unit.

CW continuous wave (laser).

D

DAC digital-to-analog converter.

DBQW double-barrier quantum well.

DBR Distributed Bragg reflector.

DC direct current.

DFB Distributed feedback (cavity).

DMD digital micro-mirror device.

DMUX demultiplexer.

DNN deep neural network.

DPU diffractive processing unit.

DtS digital-to-spike (encoding).

E

E/O electronic-optical.

EAM electro-absorption modulator.

EDFA erbium-doped optical fibre amplifier.

EEL edge-emitting laser.

F

FC/APC angled physical connector.
FC/PC physical-contact connector.
FPGA field-programmable gate array.
FWHM full width (at) half-maximum.

G

GC grating coupler.
GPU graphics processing unit.
GS grayscale.
GSG ground-signal-ground (pin layout).
GST germanium-antimony-tellurium phase-change material.

H

H-H Hodgkin-Huxley model.

I

IOP Institute of Photonics.
ISO optical isolator.
IVP initial value problem.

J

JIT just-in-time (execution).

L

LD laser diode.
LED light-emitting diode.
LIF leaky integrate-and-fire.
LLM large language model.

M

MAC multiply-and-accumulate (operation).
MBE molecular beam epitaxy.
ML Machine learning.
MQW multiple quantum well.
MRR micro-ring resonator.
MZI Mach-Zehnder interferometer.
MZM Mach-Zehnder modulator.

N

NDR negative differential resistance.

O

O/E opto-electronic.
O/E/O optical-electronic-optical.
ODE ordinary differential equation.
ONN optical neural network.
OSA optical spectrum analyser.

P

PAM pulse amplitude modulation.
PC polarisation controller.
PCM phase-change material.
PD photodetector.
PDR positive differential resistance.
PIC photonic integrated circuit.
PM power meter.
PRL photodetector-RTD-laser node.
pRL photodetecting RTD-laser node.

PRL-MN photodetector-RTD-laser monolithic node.	SMA SubMiniature version A (connector).
pRTD photodetecting resonant tunnelling diode.	SMF single-mode optical fibre.
pSNN photonic spiking neural network.	SNIC saddle-node on invariant circle (bifurcation).
PVCR peak-to-valley current ratio.	SNN spiking neural network.
	SNR signal-to-noise ratio.
Q	SOA semiconductor optical amplifier.
	SRL semiconductor ring laser.
QD quantum dot.	STDP spike-timing-dependent plasticity.
QW quantum well.	SVD singular value decomposition.
R	T
	TCGPU tensor core graphics processing unit.
R&F resonate-and-fire.	TDM time-division multiplexing.
ReLU rectified linear unit (activation function).	TIA trans-impedance amplifier.
RF radio frequency.	V
RGB red-green-blue.	VCO voltage-controlled oscillator.
RNN recurrent neural network.	VCSEL vertical cavity surface emitting laser.
RT OSC real-time oscilloscope.	VLSI very large scale integration.
RTD resonant tunnelling diode.	VOA variable optical attenuator.
RTD-SpikeSLP RTD-powered spiking single-layer perceptron.	W
RZ return-to-zero (encoding).	WDM wavelength-division multiplexing.
S	X
	XOR exclusive OR (logic gate).
SiPh silicon photonics.	
SL semiconductor laser.	
SLM spatial light modulator.	
SLP single-layer perceptron.	

Acknowledgements

First of all, I would like to extend my deep thanks and gratitude to my primary supervisor, Dr Antonio Hurtado. During our time working together, he has created a dynamic, motivating, encouraging environment that allowed me to learn and flourish to the fullest possible extent. He has always provided valuable guidance, and served as the best mentor a graduate student could ask for. I would also like to thank my second supervisor, Prof Michael Strain, for his kind support and thought-provoking discussions regarding my research.

I would like to acknowledge and thank my fellow research group members. In particular, I would like to Dr Joshua Robertson, who was always around to help with VCSEL measurements (as well as to have a good chat and laugh). I would also like to thank to Dr Weikang Zhang, for being a good friend and my trusted RTD measurements ally, and to Dafydd Owen-Newns, for help with lab automation and our always enjoyable chats and (neuromorphic) brain-storming sessions. Additional thanks also go to our former lab members, Dr Juan Arturo Alanis and Dr Julián Bueno.

The PhD journey is a long and demanding one, and there were many people at the Institute of Photonics (IOP) that made this journey into an enjoyable one. Here, I would like to thank Dimitars Jevtics, for being a good friend, for sharing his extensive knowledge of heterogenous integration methods and for our long chats during strolls from Glasgow city to Dennistoun. I'd also like to thank Eleni Margariti, for always being up to discuss the latest (mis)fortunes of our PhD-ing and lives over a pint, and to Daniel Maclure, it was always a great pleasure to chat. I must also not forget friends more on the more SCOPE-adjacent side of things, in particular Sean Bommer and Alex Blanchard, as well as Emma Le Francois, Emma Butt and Miles Toon. I would also like

to thank our IOP administrative team, Sharon Kelly and Lorraine Annand, for being always very kind and helpful. With your help, dealing with admin matters has been a breeze and never felt like a chore. This also includes Leanore Ferrans, thank you for always kindly dealing with my (sometimes questionable) IT-related requests and ideas (such as my proposition to run an Apache reverse proxy on our lab machine).

I also have to thank our extensive network of collaborators, in particular Bruno Romeira (INL Portugal), Julien Javaloyes and Ignacio Ortega-Piwonka (University of Balearic Islands, Mallorca), José M. L. Figueiredo (University of Lisbon), and Ekaterina Malysheva and Victor Calzadilla (Eindhoven University of Technology). I would also like to kindly thank the team at Princeton University, mainly Prof Paul R. Prucnal and his team members Eli A. Doris, Simon Bilodeau, Hsuan-Tung Peng and Thomas Ferreira-Lima, for providing the exciting opportunity to conduct research as a visitor at Princeton.

My thanks also go to Dr Gergely Simon, my very good friend and flatmate in Dennistoun, for always being up to discuss and brainstorm technical (and non-technical) ideas, to watch the latest episode of *The Boys* while enjoying a pint, and for always being up to conquer another boggy Scottish Munro. Finally, I want to extend my deepest thanks to my family. My life journey has gradually unfolded through various institutes and different countries all over Europe. Despite that, you have always fully supported me and encouraged me to pursue what excites me. Without all your help and support, I would not be where I am (and who I am) today.

Chapter 1

Introduction

THIS chapter provides the introductory remarks and research motivation, as well as an overview of key background concepts related to the topics and devices studied within this thesis. Furthermore, it also introduces all the components and laboratory equipment used for the experiments carried out in this thesis, and provides an outline of the following chapters.

1.1 Motivation

We are living in an age of information, with the magnitude of data production (particularly in the case of unstructured data) growing exponentially [1]. However, there is a gap between raw *data* and *knowledge*, as data needs to be processed into useful and relevant insights. According to the type, size and structuring of the data, this task can be anywhere between trivial to extremely challenging. The rapidly growing fields of data science and big data both focus on this particular fundamental challenge: extracting useful insights and knowledge from such extensive datasets.

For a wide range of data processing tasks, our digital computers were inferior in performance when compared to humans. This is where the field of artificial intelligence (AI) and its sub-field of machine learning (ML) come into the frame. The field of AI focuses on the study of agents that process information to form plans and make decisions in order to achieve their objectives. AI is a multidisciplinary field, rooted on the foundations of diverse disciplines including mathematics, statistics and logic,

computer science, linguistics, neuroscience, and decision theory. ML then focuses on giving machines the ability to "learn", that is to adapt and operate without requiring exact instructions from the operator. For example, differentiating a cat from a dog on an image is a simple task even for a child, but remained particularly difficult for the computer prior to adoption of artificial (in this case convolutional) neural networks and the related training procedures.

It is not an overstatement to say that we are currently living in the golden era of AI [2]. Following a series of previous boom-and-bust cycles (waves of excitement in "intelligent machines" followed by the so-called "AI winters"), we are currently (since approx. 2010) witnessing an astonishing growth in the field and its permeation into basically every domain of our society. While commodities such as oil are still of utmost importance for mankind, some people view data as the most valuable commodity of a modern society. In turn, any human activity which either collects or processes information (data) can directly benefit from smarter, more advanced data processing techniques and approaches, among which AI and ML are the most promising and powerful. Therefore, both AI and ML have a tremendous potential to influence and empower virtually every existing domain of human expertise. This renders AI as arguably the most disruptive technology of our times, with potential to synergistically assist people in their work (or perhaps even completely replace them). Furthermore, in order to enable modern technologies such as self-driving cars, personalised machine-assisted medical care or reliable natural language processing, advanced AI approaches represent the key building block.

1.2 Overview of key concepts

1.2.1 Artificial neural networks (ANNs)

The artificial neural network (ANN) is one of the key concepts underpinning the fields of ML and AI. An ANN is a graph based on abstraction of the architecture of a biological brain. In this graph, the network nodes are called neurons or cells, and the links (weighted connections) are called synapses. Many different ANN architectures exist,

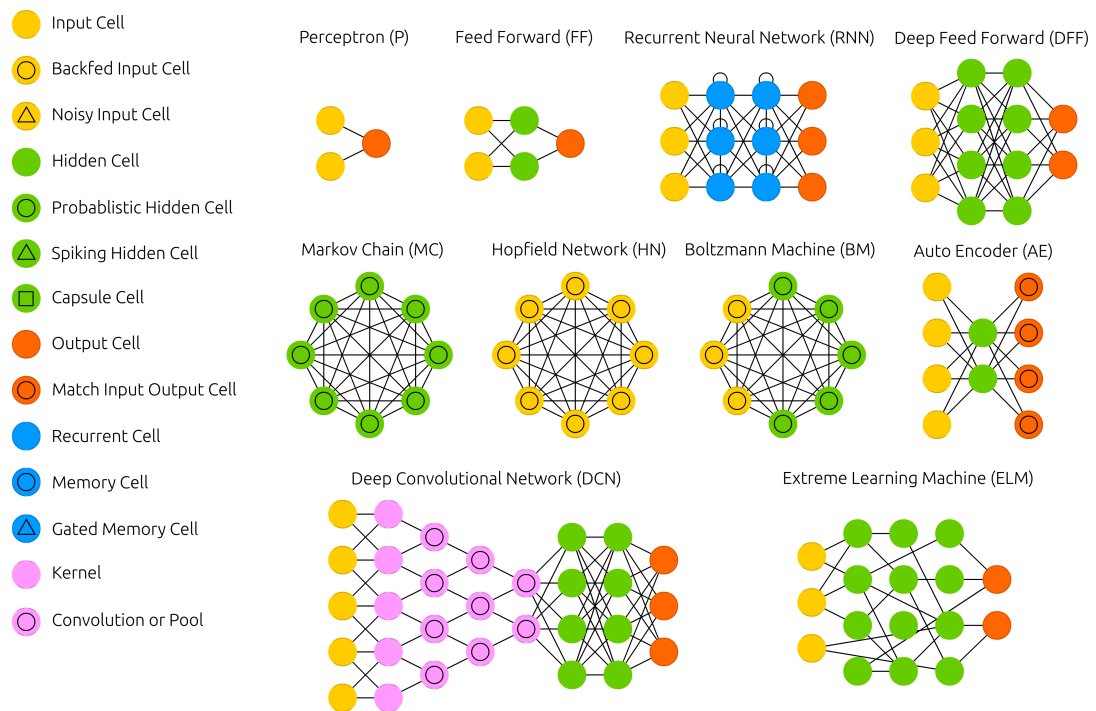


Figure 1.1: Excerpt of "The Neural Network Zoo", demonstrating various architectures of artificial neural networks. Reproduced and modified from [3].

covering a range of various use-cases. Few examples of ANN architectures are shown in the excerpt from the "Neural Network Zoo" [3] included in Fig. 1.1.

Artificial neural network architectures can be classified as belonging to distinct 'generations' [4]. In the first generation networks, the computational units (neurons) are based on the McCulloch and Pitts models, and utilise a linear combination of input signals followed by a step-like activation function, producing a binary output. These networks are boolean complete [5]. Examples of such networks include perceptrons, multi-layer perceptrons and Hopfield networks [4].

The second generation of neural networks replaced the step activation function with a continuous activation function (such as the sigmoid function), therefore producing continuous output values. In the simplest case of feed-forward neural networks, information is transformed from the input to the output. The neurons are typically arranged in layers, with arrangements of more than two (input, output) layers being referred to as deep neural networks (DNNs). Such networks effectively operate as

function approximators via composition of differentiable functions. DNNs allow to express complex data-label relationships by stacking the nonlinear transforms [6], and are typically trained using backpropagation. The error backpropagation algorithm, first proposed by Seppo Linnainmaa in 1970 [7] and later demonstrated as being beneficial for neural network training by Rumelhart in 1986 [8], minimises network’s loss-function using gradient based rules such as automatic differentiation.

Here, it is worth noting that ANNs are algorithms run on digital computers and represent a very coarse abstraction of actual biological neural networks. Functionality wise, ANNs do not bear close resemblance to how the computational primitives in the brain operate. A case can be made that different environments (digital operations in silicon computers vs. evolution-shaped biochemical procedures in living tissues) prove different resources and constraints [9], and therefore forcing principles from biology into the digital domain may not be directly associated with better algorithmic performance [10]. Nonetheless, the brain remains a source of inspiration for designing more efficient AI algorithms.

1.2.2 Spiking neural networks (SNNs)

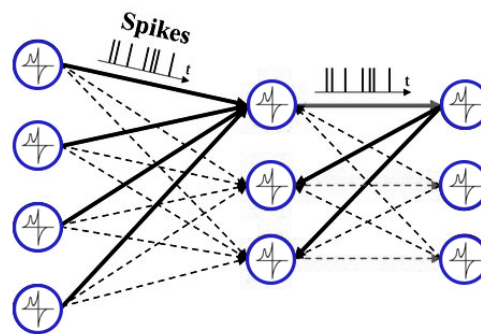


Figure 1.2: Simplified schematic of a general spiking neural network.

In ANNs, network nodes are typically continuous function approximators with a given activation function (ReLU, sigmoid and many others) operated on a digital, clocked system. Meanwhile, biological neurons operate and signal using asynchronous events: action potentials (spikes) that can be considered as digital in their amplitude,

yet analog (temporally precise) in time [11]. Such combined analog-digital computation is believed to represent a very resource efficient form of computation [12]. The chasm between ANNs and biology is crossed by artificial spiking neural networks (SNNs, Fig. 1.2), which aim to operate using event-based signalling like biological neurons, and are sometimes referred to as the third generation of neural networks. Studies have shown that SNNs (with temporal spike coding) exhibit at least the same computational power as previous neural network generations, while requiring significantly fewer neurons for comparable performance in certain tasks [4].

Training methods for spiking neural networks (SNNs) typically differ from those employed for training of non-spiking, conventional ANNs. Unfortunately, traditional backpropagation does not readily yield itself to be used in training of SNNs due to the non-differentiable character of the spike generation activation function [13]. Nonetheless, extensions and variations of backpropagation have been introduced specially for training of SNNs. These include backpropagation-through-time (BPTT), which borrows concepts from methods used to train recurrent neural networks (RNNs), the SpikeProp [14] algorithm, backpropagation with derivative approximation for LIF-class neurons with demonstration in sparse, deep convolutional SNNs [13] or the SLAYER (spike layer error reassignment) algorithm [15].

However, restricting focus to backpropagation modifications might not provide the optimal method for SNN training. Because these networks employ time during computation and can utilise biologically-plausible local learning rules (such as spike-timing dependent plasticity (STDP) and long-term potentiation, among others), these new functional primitives may hold the key for efficient training of these networks. STDP represents a specific kind of Hebbian learning approach which is believed to constitute part of the learning process in networks of biological neurons. There is already a wide (and growing) variety of algorithms specially designed for SNN training. Examples include ReSuMe [16], SuperSpike [17] (which omits temporal dependency) and event-based EvSpikeProp and EvSpikePropRT [18], where the derivative term is defined from firing time. Furthermore, spike-timing and temporal domain encoding can be used with backpropagation-like algorithms to achieve training of SNNs, as

shown in the ANN-conversion-aided procedure in [19], and backpropagation-like credit assessment was reported as indirectly performed in SNNs using energy-minimisation [20]. With respect to neuromorphic computing, not every learning algorithm yields itself to direct implementation in hardware, as hardware engineering is ruled by more significant constraints than software engineering. One of the promising recent examples of a hardware-viable training method is EqSpike [21]. EqSpike is an equilibrium propagation in spike-based networks which offers a locally-computed alternative to conventional backpropagation and which has been numerically demonstrated to provide MNIST benchmark dataset performance comparable to state-of-the-art BPTT method.

1.2.3 Nonlinear dynamics and excitability

The field of *dynamics* deals with systems that change and evolve in time. Such dynamical evolution can have many different shapes and forms: a system can be slowly reaching a stable equilibrium, it can repeatedly oscillate, or behave in a much more complex way, including chaos. Typically, this dynamical evolution of a system (in continuous time) is captured using differential equations. These systems can be broadly divided into two classes: linear or nonlinear. In linear dynamical systems, the differential equations are linear - the right-hand side of the dynamical differential equations is completely described as a weighted sum of the state variables: $\frac{\delta x}{\delta t} = Ax(+B)$. If the equation contains more complex terms (such as products, powers, and functions), the system it describes is considered nonlinear. Nonlinear systems are much more interesting as they can describe very complex dynamical behaviour, yet most are unfortunately also impossible to solve analytically [22]. The devices studied in this thesis represent examples of nonlinear systems.

Interestingly, nonlinear systems (accompanied by a controller) can be regarded as universal computing machines [23]. Since these systems can be perturbed and influenced by external signals to which they react, they must (in some way) process information [24], with processing (mapping) functionality varying between different nonlinear systems [25]. Perhaps the most striking demonstration of this functionality is in one particular class of physical "computers" called reservoirs, where inner nonlinearities

are used for information processing with a single, trainable output layer. Even a system such as a soft robotics tentacle suspended in water tank has been shown as suitable for reservoir-based computation [26]. Reservoir computers represent a very interesting class of systems, and there are ongoing research efforts within our research group at University of Strathclyde on photonic reservoirs (see for instance [27]).

An important concept in study of dynamical systems are bifurcations. A bifurcation represents a points where a dynamical system undergoes a major change in its dynamical behaviour, when a specific parameter crosses a critical value. Saddle-node [28] is a type of dynamical bifurcation during which two system equilibria (a stable and an unstable fixed point) collide and annihilate. Saddle-node bifurcation is one of the bifurcations closely tied to the concept of excitability [29].

Excitability refers to a particular dynamical behaviour in nonlinear systems, where a perturbation above a given threshold results in well-defined and pronounced response, whereas sub-threshold perturbations do not elicit any significant response and keep the system in a steady state. This may take form of voltage spikes (action potentials) in biological neurons and heart muscles, calcium waves in cells [30], chemical concentration spikes in reaction-diffusion systems, or light intensity spikes in lasers. For illustrative purposes, the toilet flushing mechanism can also be considered as an example of an excitable system.

Some characteristics of excitable behaviour are:

- From a dynamical standpoint, the system has a single stable attractor (resting/equilibrium state) [31] at which the system rests when unperturbed.
- Perturbation or stimuli below a given threshold induce negligible response in the system.
- Perturbation or stimuli above a given threshold trigger a pronounced reaction from the equilibrium state. Once this response is elicited, the system undergoes a well defined, large excursion in the phase space before settling down.
- The character (phase trace) of an excitable system response is independent on the stimulus.

- After the excitable excursion, the system settles back to the attractor and in this period, it does not respond to new perturbations even if these are super-threshold [32]. This temporal interval is usually referred to as refractory (or sometimes lethargic) period.

One of the original works on excitability in semiconductor lasers was presented by [33], where the refractory period is, at the material level, governed by recovery of carriers in the semiconductor. The theoretical framework for understanding of excitable phenomena in lasers was also provided by [32].

1.2.4 Dynamics of biological neurons

The aim of understanding and modelling behaviour of neuronal cells represents one of the important research directions in the field of neuroscience [34]. Some of these models aim to capture the neuronal behaviour as closely as possible, in its full dynamical extent. These include the classical Hodgkin-Huxley (H-H) model, which was introduced in 1952 based on measurements of action potentials from the squid's giant axon [35, 36]. This model, based on four-dimensional set of nonlinear differential equations, is complex, accurate and allows for modelling of range of biological neurons. Following that, a simplified model in the same framework was developed by Morris and Lecar in 1981 [37]. These models are based on biophysically meaningful sets of parameters and take into account the biochemical and electrochemical phenomena in biological neurons. While such approaches are ideal for modelling and understanding the brain, the complexity of these models makes their hardware implementation challenging [38]. A further simplification of the H-H model is the FitzHugh-Nagumo model [39], which is based on the van der Pol oscillator.

A second group of neuronal models represents phenomenological models that focus more on capturing the fundamental behaviour observed in neurons. Leaky integrate-and-fire (LIF) is a widely adopted spiking neuronal model used in computational neuroscience, AI and neuromorphic engineering. In LIF neurons, synaptic inputs are integrated both spatially and temporally. Each input is reflected in the change of an inner state variable, which is leaky and gradually decays over time (typically with an exponential decay).

Once the state variable reaches a given threshold, a spike is elicited. Following spike firing, the state resets and the system enters a refractory period. It has been shown that networks of leaky integrate-and-fire neurons are capable of simulating both Turing machines and traditional neural networks [4]. A different type of dynamical behaviours is captured in a resonate-and-fire (R&F) model [40], where spike firing is more sensitive to frequency (rate) of input perturbations. R&F and LIF mechanism form the basis for classification of neurons into Type I (integrator) and Type II (resonator) neurons [41].

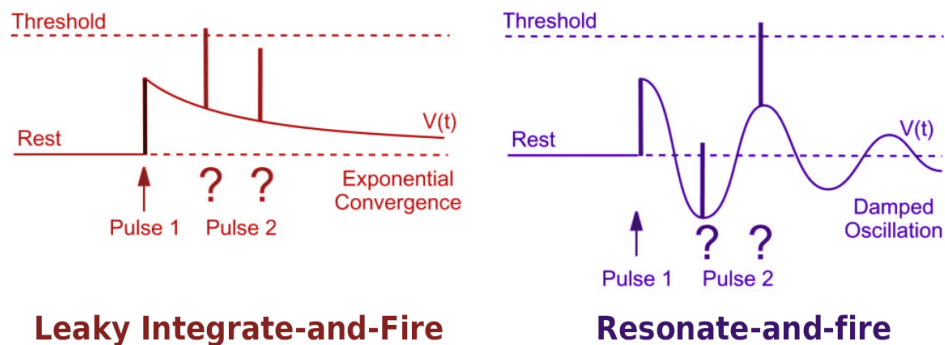


Figure 1.3: Comparison of spike firing mechanisms in integrate-and-fire and resonate-and-fire neuronal dynamical models. Adapted from [40].

Based on the character of produced spiking responses, biological neurons can also be separated into various classes. In *Class I* neurons: [42], action potentials (spiking) can be generated at wide range of frequencies that depend on the strength of the applied perturbation (current). Dynamically, this is enabled by a saddle-node bifurcation. This dynamical behaviour can be observed for example in cortical excitatory pyramidal neurons [34]. In *Class II* neurons, a sequence of action potentials is generated in a specific range (band) of frequencies, with lower influence of perturbation strength. These types of "resonator" neurons include brainstem mesencephalic V neurons, stellate neurons of the entorhinal cortex and cortical inhibitory interneurons [34]. Conventionally, these neurons are described by models exhibiting a Hopf bifurcation, at which the stable point changes to a limit cycle [43]. Once a spike is emitted, the system relaxes back to the stable fixed point via relaxation oscillations, which allows for these resonance effect to arise. The Fitzhugh-Nagumo model is among such models exhibiting a Hopf bifurcation. Finally, in *Class III* neurons, the resting state remains stable for any

biophysically-feasible input current, and spikes are fired only at the transient of the perturbation [44].

1.2.5 Semiconductor lasers

Lasers represent a class of active devices that operate as light sources. Importantly, the light generated by lasers is collimated (highly directional), coherent (maintaining phase across space and time) and monochromatic (the emission wavelength range is typically very narrow). A laser requires at least two key functional parts for its functionality: an active (gain) medium and an optical resonator, which feeds the light back into the active medium via reflectors. The key process for lasing is stimulated emission, a process which generates coherent "duplicates" of input photons in the active medium. By ensuring a specific energy distribution of charges in the gain medium (population inversion) that increases the likelihood of stimulated emission, and by feeding the photons through the medium many times as standing waves in the resonator, net photon gain per each pass can be achieved in the system, resulting in lasing.

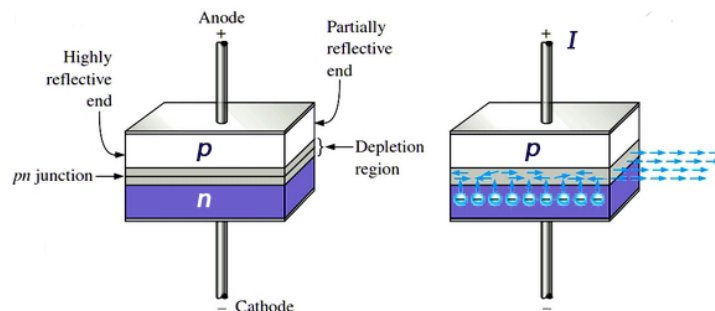


Figure 1.4: Simplified schematic of a semiconductor laser (laser diode) with a Fabry-Perot type cavity. Reproduced and adapted from [45].

There are various technologies that can be used for the gain medium in lasers. Based on the active medium, lasers can be classified as gas lasers, liquid lasers, semiconductor lasers, solid-state lasers and fibre lasers. In particular, semiconductor lasers (SLs) are a cheap and compact-sized class of lasers that are available over a wide range of wavelengths. Their working principle is similar to that of a forward-biased p - n junction based on direct-bandgap semiconductors, and is shown in a simplified schematic in

Fig. 1.4. At this junction between the p -doped and n -doped materials is the active region, where electrons are injected from the n -layer and holes are injected from the p -layer when under forward biasing. Recombination of these confined excitons produces photons by the process of *charged carrier radiative recombination*, with energy roughly equal to that of the bandgap. Under population inversion, the likelihood of stimulated emission is higher than photon absorption, and lasing can be achieved. Pumping of SLs is typically achieved via electrical current injection. The resonator cavity in SLs is typically realised using Fabry-Perot cavities, ring cavities, distributed feedback (DFB) cavities or with Distributed Bragg reflectors (DBRs) [46]. Furthermore, single or multiple quantum well (QW) heterostructures are frequently incorporated in the active regions of modern SLs to enhance quantum efficiency, modulation speeds and lower the threshold current and spectral linewidth. The choice of material system influences the band structure and bandgap, which in turn governs the emission wavelength and, among other properties, the spectral linewidth enhancement factor α , an important parameter prominent in SLs. This α factor describes phase change arising from variation in medium refractive index n as a function of carrier density [47] and increases the linewidth of SLs. Furthermore, it also influences the dynamics of SLs [48]. The vertical cavity surface emitting laser (VCSEL) used experimentally in Chapters 3, 4 represents a case of a SL.

1.3 Description of the components and equipment used in the experiments

Throughout this work, various photonic and optoelectronic systems have been experimentally investigated. In the vast majority of cases, these experiments were realised using standard, telecom-wavelength, off-the-shelf fibre-optic components. In addition, radio frequency (RF) and generic electronic probing equipment has been used. This Section will provide an overview of all of these components and laboratory equipment used during the experiments.

1.3.1 Laser driving components

- **Laser diode mounts:** *Thorlabs LDM21, LDM90*. These are commercially available standard laser diode mounts with embedded precision temperature control for laser stabilisation and operational conditions control. Additionally, the Thorlabs LDM90 includes a built-in bias-tee network that allows for RF signal modulation of the laser current. The 50Ω input is AC-coupled directly to the laser with upper bandwidth limit provided by the manufacturer at 600 MHz.
- **Laser temperature controllers:** *Thorlabs TED200C*. These controllers directly connect to the laser mounts and allow both for monitoring and control of the temperature of the thermistor, with built-in feedback mechanism for maintaining the set temperature conditions. Unless specific emission wavelength is required, the lasers are typically stabilised at room temperatures (293 K).
- **Laser drivers:** *Thorlabs LDC201CU, Thorlabs LDC205C*. These controllers directly connect to the laser mounts through a proprietary interface and allow for high precision (10 μ A resolution) laser (VCSEL) biasing.

1.3.2 Fibre-optic components

- **Optical isolator (ISO)**. Allows for single direction of light transmission only, therefore preventing unwanted back-reflections (via high attenuation (>40 dB) of backward propagating optical signals). This improves signal-to-noise ratio in readouts, protects components in the setup and prevents unwanted feedback.
- **Variable optical attenuator (VOA)**. Allows for controllable optical power adjustment via varying the coupling efficiency to the output fibre. Power control is essential to establish well-controlled experimental conditions and to protect more sensitive components and equipment.
- **Optical splitters/couplers**. These components allow to split and combine optical signals with a given power ratio. Various coupling and port arrangements

have been used in this work, including 50:50 and 90:10 four-port couplers, and a 25:25:25:25 1-to-4 splitter.

- **Polarisation controllers (PC).** These components utilise the mechanical stress-induced optical birefringence effect in optical single-mode fibre. The effect is controllably introduced in a fibre loop, created by wrapping the fibre inside multiple (typically three) paddles. Manual adjustments of these paddles allows for output light polarisation state control through twisting of the fiber, which causes stress-induced birefringence.
- **Optical circulators (CIRC).** These are three-port fibre-optic components that only allow light to directionally pass between selected pairs of ports while introducing significant losses (≈ 50 dB) in the remaining directions. Typically, the single-direction flow of optical signals is enabled between ports $1 \rightarrow 2$, and between ports $2 \rightarrow 3$. Any input to port 3 experiences large attenuation. This functionality is achieved through the magneto-optic Faraday rotation effect and birefringence. Within this thesis, circulators are typically used to achieve both signal injection to and collection from VCSELs.
- **Single-mode fibres (SMF).** Used to connect all the optical components forming the experimental setups. All the fibres used are non-polarisation maintaining (non-PM).
- **Optical power meter (PM):** *Thorlabs PM20C*. This handheld optical power meter utilises an InGaAs detector and operates in the wavelength range 800 nm-1700 nm. The investigated wavelength can be readily adjusted by the user. Mean power measurements (sampling rate 10 Hz) are available in the 1 nW to 20 μ W. In typical measurements, the power would be taken directly from the main signal path, or a splitter (50:50 or 90:10) is used for continuous power readout.
- **Erbium-doped fibre optic amplifier (EDFA).** EDFAs are optical devices available in a specific wavelength range (telecom C-band, around 1550 nm) that allow for amplification of optical signals without relying on O/E and E/O conversion.

Instead, they utilise a section of optically pumped rare-earth-doped fibre (with erbium ions) that provides input laser light amplification via stimulated emission. The erbium-doped fibre is typically pumped using one or multiple laser diodes. The EDFA model used in our experimental work was the Thorlabs EDFA100S (core-pumped, single mode, non-polarisation maintaining, > 20 dBm output power).

1.3.3 Electronics equipment used in the experimental work of this thesis

- **Arbitrary waveform generator (AWG):** *Keysight M8190A* (12 bit @ 12 GSa/s, 2x channels). This machine represents a key piece of experimental equipment that has been used as a source of all RF signals used to investigate devices within this thesis. The AWG is fully programmable, with support for as low as 83 ps sample length and up to 600 mV peak-peak in direct sourcing range, up to 1000 mV peak-peak in DC amplified sourcing range. To control this multi-channel source of RF signals, a custom Python wrapper was designed based on the `pymeasure` library for direct control and waveform handling. These wrapped functions are part of the lab toolkit developed by the author of this thesis*.
- **Real-time oscilloscopes:** *Agilent Infiniium DSO81304B* (40 GSa/s sample rate, 13 GHz bandwidth), *Rohde & Schwarz RTP* (40 GSa/s sample rate, 16 GHz bandwidth). The real-time oscilloscope allows for acquisition and analysis of dynamical (RF) signals produced by all the devices under investigation. The Agilent Infiniium oscilloscope was used for some of the first measurements with VCSELs, and was later superseded by the RTP oscilloscope which offered higher bandwidth and memory. Custom acquisition modules for remote operation were written for the RTP to allow for increased convenience of lab measurements, based on the `rsinstrument` Python library[†]. In addition, single-shot, multi-channel

*<https://github.com/mhejda/StrathLab>

[†]<https://github.com/Rohde-Schwarz/RsInstrument>

data acquisition functionality was implemented from scratch using SCPI and VISA/pyvisa protocols[‡].

- **Amplified photodetectors: Thorlabs PDA8GS.** Photodetectors were required for conversion of optical signals into the electrical domain for analysis, and for investigation of the direct optical input functionality in the RTDs (Chapter 4). Due to the low-power nature of the signals from the VCSELs, amplification is typically required. The PDA8GS is a Fixed Gain InGaAs Amplified Photodetector operating at 750 nm - 1650 nm wavelength range, with bandwidth ranging from DC to 9.5 GHz.
- **Precision sourcemeter: Keithley 2450.** A high precision, high sensitivity (20 mV, 10 nA) benchtop DC source with direct response acquisition and programmable sweep programs was used to measure the current-voltage (I - V) characteristics of the RTDs under test. This system was also the primary DC source used for probing and characterising RTDs (Chapter 4). This unit is programmable, and was also incorporated into the StrathLab remote control and measurement software toolkit.
- **Stable DC sources: RS Components Pro IPS 603** (single channel 200 W Bench Power Supply), *Teledyne LeCroy T3PS43203P* (4 channel 217 W Bench Linear Power Supply). These sources were used for providing bias voltages to all the optical and electronic devices that required external power: optical modulators, RTDs, RF amplifiers etc.

1.3.4 Radio frequency (RF) probing equipment

- **RF probes: GGB Picoprobe Model 40A.** A standard microwave probe with low insertion loss. Three-tip footprint in the Ground-Signal-Ground (GSG) is used, with tip spacing (pitch) of 100 μm . Connection is provided through female K-connectors (2.92 mm, reverse compatible with SMA).

[‡]<https://github.com/mhejda/StrathLab>

- **Micropositioners:** *Signatone S-M40*. Three-axis micropositioner specifically designed for RF probing. It has 2 μm precision, 25 mm travel distance in all three axes. The positioner head includes planarity adjustment for multi-contact probing (such as GSG). Models used were straight arm (E/W) orientation and used magnetic mounting base. These were placed and secured on passivated stainless steel breadboards from Thorlabs.
- **Bias-tees:** *API-Inmet 8800SMF1-12*. The bandwidth of this wideband bias-tee is 10 MHz to 12.4 GHz. The impedance is 50 Ω , and the component is connected through SMA connectors.
- **Resistive splitters:** *Mini Circuits ZFRSC-183-S+*. This three-port device is a wideband (DC to 18 GHz), 2-way-0° resistive 50 Ω power splitter. In the experiments, this is typically used to split RF signals for analysis, or to combine three signal branches together.
- **Wideband electronic amplifiers:** Mini Circuits ZX60-14012L-S+ (10 dBm inverting, 0.3 GHz-14 GHz). Due to the wide range of operational voltages for the Mach-Zehnder modulators (used to externally modulate the CW light signals generated by the tuneable laser sources), additional amplification for the AWG RF outputs was required, and was realised using this simple amplifier. The inverting operation was compensated either software-wise or using the inverting output RF port of the AWG.
- **Interconnects:** SMA (Sub Miniature version A) and K-connectors (2.92 mm)

1.4 Thesis outline

The thesis is structured in six chapters.

Chapter 2 provides an introduction into the field of neuromorphic engineering, with particular focus on light-powered neuromorphic systems. The motivation, definition and characteristics of general neuromorphic systems are introduced, and a short summary on the currently leading approaches realised on electronic platforms is also provided.

Furthermore, the field of neuromorphic photonics is introduced and a summary of the key advantages and challenges of such systems is provided. Finally, Chapter 2 also reviews selected existing optical neural network and light-powered neuromorphic approaches based on their implementation and functionality.

Chapter 3 constitutes the first results chapter. Here, the vertical cavity surface emitting laser (VCSEL) is introduced, including an overview of the main characteristics and properties of this type of semiconductor laser diode. The methods to operate the VCSEL as an all-optical, high-speed photonic spiking neuron are introduced, and experimental characterisation of the excitable dynamics and spiking responses is provided, including the evaluation of the refractory period and demonstration of its leaky integrate-and-fire functionality. Based on these findings, experimental insights are provided on using these photonic spiking VCSEL-neurons as neuromimetic rate encoders, with capability to directly convert digital image data into temporally coded spike trains. Finally, experimental validation is provided for operation of VCSEL-neurons in conjunction with integrated silicon photonics weight banks.

Chapter 4 constitutes the second results chapter of this thesis. Here, the resonant tunnelling diode (RTD) is introduced, including its dynamical model and characteristics, and both theoretical modelling and experimental results on RTDs are provided. First, the model is introduced and the effects of parameters on the dynamical evolution of the system are discussed. Furthermore, excitable phenomena arising in these devices are studied using the model, and an extension from an electronic RTD circuit to an optoelectronic (O/E/O) model is provided. Following that, both the model and the experimental setup are used for characterisation of the excitable dynamics in these devices. Phenomena which arise from the dynamical character of the RTD (such as resonate-and-fire excitability) are also introduced and characterised.

Chapter 5 constitutes the third and final results chapter of this thesis. Based on the modelling and experimental findings from the previous chapter, the optoelectronic RTDs are utilised both in theory and experimentally to demonstrate information processing functionalities. These include tasks performed with a single node, such as experimental demonstration of coincidence detection and exclusive OR (XOR) logic gate. Numerically,

sequential convolution-like functionality for Tetris-pattern recognition and image feature extraction in a single RTD-MN node are demonstrated. Furthermore, multi-node and networked arrangements are investigated. Information propagation between two optoelectronic RTDs is demonstrated experimentally with a dynamically weighted interlink. Finally, a small-scale, perceptron-inspired optoelectronic SNN is realised in theory. The network is trained using a hybrid, supervised, spike timing dependent rule, and is used to perform spatiotemporal pulse pattern recognition.

Chapter 2

Neuromorphic photonics and optical neural networks: an overview

THIS chapter provides the background and summary of key concepts related to the rapidly growing and evolving field of neuromorphic engineering, with main focus on light-powered (photonic) implementations of such systems. Therefore, it provides the foundations for the devices and systems proposed and investigated in this thesis.

2.1 Introduction

As once said by the former CEO of Xerox Corp., Mark D. Weiser, truly paradigm-changing technologies embed themselves into the fabric of our day-to-day lives and become basically invisible, naturally extending our capabilities in a seamless manner and impacting our everyday decisions with radical ease-of-use [49]. Apple's iPhone can be considered as an example of such technology, bringing all of mankind's shared knowledge (internet), communication tools, navigation and entertainment in a device that is now ubiquitous and can be easily operated even by kids.

Artificial intelligence (AI) clearly follows a very similar trend. Nowadays, we are already witnessing the proliferation and seamless integration of AI and ML-based

approaches into our day-to-day lives. For example, artificial intelligence approaches are widely used for natural language processing, including translation, query processing in search engines [50] and filtering of junk emails and unsolicited (spam) messages [51], as well as to achieve human-like performance in environments with limited sets of conditions such as games [52]. Among the recent remarkable achievements of AI is AlphaFold II [53, 54], which achieved successful prediction of a folded (ternary) protein structure based on the provided amino-acid sequence, therefore solving a very significant and difficult problem from the field of computational biology. Notably, 2022 was the year in which natural language processing, large language models (LLMs) and generative AI moved into the general public spotlight. This was mainly due to introduction of large models such as Midjourney, DALL-E2 [55], Imagen [56] and Stable Diffusion [57] that allow for on-demand, prompt-based synthesis of arbitrary photorealistic or abstract imagery. Furthermore, after previously introducing the record-breaking third generation of the Generative Pre-Trained Transformer (GPT-3) [58] that boasted 175B parameters, OpenAI publicly launched the new model called ChatGPT in December of 2022. Based on reinforcement learning from human feedback (RLHF), the publicly available, prompt-based AI chatbot shocked the world with its capability to deliver meaningful written text indistinguishable from human writers and demonstrating capabilities such as writing computer code [59]. The disruptive potential of ChatGPT is hard to overstate [60], and goes far beyond the discussed scope. Just prior to submission of this thesis in June 2023, GPT-4 by OpenAI was released and demonstrated astonishing performance in range of very complex and open-ended tasks, including passing (and outperforming humans) in standardised professional exams such as the American Uniform Bar Examination [61] required in the US for law practitioners. Furthermore, the potential for significant augmentation of human activities is not limited to languages, as AI-powered models can also enable autonomous transportation or advanced, computer-aided medical diagnostics.

While LLMs and ultra-large scale AI models deliver impressive functionalities, these are not the only types of AI systems that can provide significant utility. On the other end of the scaling wars are edge-computing approaches, which strive to trade off the broad scope of functionalities of large models for very low power consumption at a

specific task or functionality. As an example, a dedicated chip with a small-scale deep neural network (DNN) is implemented in a low-power hardware for always-on voice activation of personal voice assistant functionality on Apple devices [62]. The "Hey Siri" phrase recognition is performed by the co-processor and serves as a detector. When the likelihood score from the always-on chip exceeds a threshold, the co-processor requests analysis of the recorded phrase using a larger, more precise DNN in the main processor. Without the co-processor, the higher precision DNN operating in always-on mode would quickly drain the device battery. Furthermore, edge ANNs and deep learning approaches can be used for processing of microphone, gyroscope, accelerometer and light sensor data in portable devices, allowing for biometric verification or user emotion recognition [63]. Research and activities in this particular type of AI and edge processing are sometimes labelled as tinyML and driven by, among others, the tinyML Foundation*.

2.1.1 Motivation for alternative computing approaches

The previous section highlighted some of the notable representatives of the recently introduced, powerful AI models. Nowadays, these models bring significant, practical utility to our society, and as such, we are likely to see them employed in growing number of use cases and sectors. Whilst we are nowadays seeing staggering advances in computer science and significant media coverage of AI, the hardware enabling operation of these receives in comparison less attention. In this section, I will briefly cover some of the current challenges present in AI, as well as some of the motivations justifying the search for new generation of AI-specific hardware.

Resource requirements of state-of-the-art AI models

While new AI algorithms seemingly blur the boundaries between the creative efforts of humans and outputs of machines, they carry a significant (albeit somewhat hidden) cost that will only get more important as the field grows. Yann LeCun, one of the most prominent AI researchers, once said that *"One interesting property of [deep learning] systems is that the larger we make them, the better they seem to work."* [64]. This is

*<https://www.tinyml.org/>

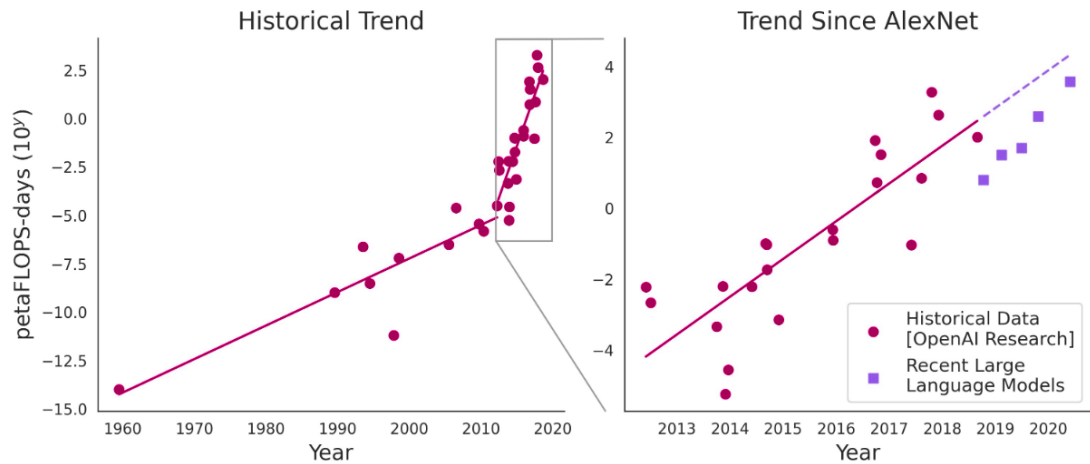


Figure 2.1: Illustration of growth in compute demands for AI models. During the 'AI golden era' of the past decade, rate of growth in compute requirements of new models has significantly outpace the historical norm. Reproduced from [65].

demonstrated by the previously mentioned LLM GPT-3, which is a model employing 175B parameters [58]. This model was released in May 2020 and was over a magnitude larger than the previous largest model, Microsoft's Turing NLG from February 2020. This highlights the astonishing pace at which these models are growing in scale and performance. A figure showing the evolution of computational requirements for AI models is shown in Fig. 2.1. Since the mid-2010s, a rapid upscaling trend can be observed where the computational requirements of AI models double approximately every 3.5 months [65].

To provide a better framing of the computing requirements of current models, training the GPT-3 model requires approx. 3.14×10^{23} FLOPS (floating-point operations) [66]. Assuming a performance of 130 TeraFLOPS for a single, AI tailor-made NVIDIA V100 tensor core graphics processing unit (TCGPU), a simplified approximate figure for one training procedure for the GPT-3 model stands at 28 days utilising one thousand V100 TCGPUs running in parallel [67]. At the time of the writing of this thesis, the price of one such TCGPU is quoted at ≈ 2000 US dollars, hence setting the cost at two million dollars solely for the required AI accelerators that allow one to achieve training of the model in a feasible amount of time, all without inclusion of costs for any other hardware or electricity overheads. This approximate figure is only for illustrative purposes and

does not account for practical challenges such as computation parallelisation, model memory requirements etc.

In 2019, a report by Strubell et al. [68] estimated that the training procedure for a single, transformer-powered NLP model with 213 million parameters [69] consumes 656 MWh (efficiency adjusted with multiplier of $\times 1.58$ [68]) and produces a carbon footprint of 285 tonnes of CO₂. This figure is equivalent to the carbon footprint produced on average by five petrol cars over their entire lifetime. Similarly, the Carbontracker tool [67] estimates the power consumption for training a GPT-3 model on Nvidia GPUs as 190 MWh (efficiency adjusted $\times 1.125$), yielding a total CO₂ footprint equivalent to that of a typical new European car driving for 700,000 kms [70]. Pairing these numbers together with the estimated rate of growth shown in Fig. 2.1, it is easy to see how AI's energy consumption and carbon footprint could easily become highly problematic. Furthermore, besides advances in the most powerful models, growth is also expected in the amount of smaller (edge) models deployed in the next-generation of smart personal devices, cars and healthcare instruments etc. Altogether, the sky-rocketing scaling trend of AI models, the increasing resources requirements, growing energy consumption and carbon footprint together with widespread and growing usage of these models all make a significant case for development of AI-optimized, application specific computing hardware.

Energy efficiency of the biological brain

The brain remains the most efficient supercomputer known to us. To give a comparison with the energy consumption figures reported previously for large AI models, it takes approx. 6 years for a human brain to consume 3.6 GJ (1,000 kWh) of energy (estimated from Adenosine triphosphate (ATP) consumption) [71]. While it is very difficult to draw direct comparisons between the process of thinking in the brain and computing a LLM (including the fact that there is no distinct "training phase" for the brain), the difference between the energy consumed by the brain (performing a wide variety of thinking and survival tasks, using 10^{15} synapses and 10^{11} neurons, requiring 0.16 MWh per year) and the limited-purpose digital neural network training (values approaching or even exceeding

200 MWh per training, as discussed above for transformer models) is more than striking. Interestingly, the transformer architecture (powering some of the state-of-the-art LLMs) has recently been demonstrated as computable with astrocyte networks [72], further highlighting the potential for using biologically-educated architectures for computing.

Slowing down (plateaus) in empirical transistor laws

Numerous empirical laws have effectively anticipated and forecasted developments in computer hardware and performance for many years. Moore’s law, Dennard’s scaling [73], and Koomey’s law [74] represent such empirical trends. However, these trends are slowly plateauing. Dennard’s scaling became unsustainable in approx. 2004 due to leakage currents [75], which signalled the first challenges in terms of further downscaling of transistors. Currently, the CMOS lithography feature size available in the market is around the 5 nm mark (Apple’s M1 Max and Pro chips use TSMC N5 5 nm process), with Apple and Intel expected to reach towards the 3 nm process in the near future. However, making transistors smaller no longer provides a guaranteed advantage in performance or efficiency [76], as highlighted by the generalized trends shown in Fig. 2.2. Further downscaling is also expected to be gradually more and more challenging, ultimately being limited by physical constraints. Therefore, there is a good likelihood that R&D focus will shift towards alternative computing architectures that might offer more favourable performance and efficiency improvements per dollar invested. In particular, non-von Neumann and in-memory computing approaches are considered as the viable next step ahead from the ”post-Moore age” [75].

Rise of application-specific accelerators

Modern digital computers represent high performance, versatile machines that benefit from decades of CMOS/silicon technology development. Powered by the von Neumann architecture with distinct memory and arithmetic & logic units, they are very flexible, utilise Boolean logic, bit-precise digital representation of information and operate with clocked signals and time-multiplexing. At low levels, the von Neumann architecture suffers from a bottleneck due to information and instructions sharing the same bus and

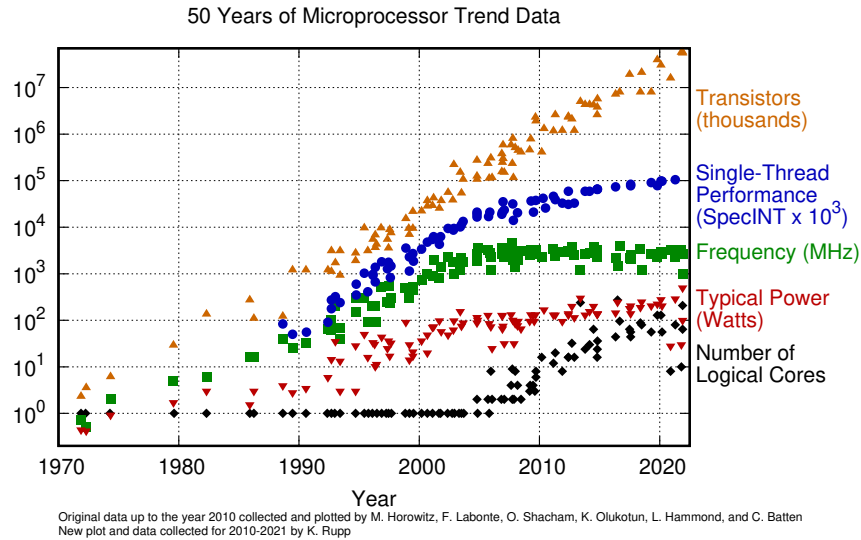


Figure 2.2: Selected key performance metrics of current digital computers, demonstrating stalling trends in performance. Available from [77].

due to these being processed sequentially, one-by-one. Furthermore, digital computers are sub-optimal for certain classes of tasks (such as optimization) or some processes (such as providing true randomness). This provides the ground for specialized, application-specific co-processors, technologies and approaches that can extend the usefulness of all-purpose digital computers. This task is already picked by graphical processing units (GPUs, which are better than CPUs for some of the algebraic operation underlying the operation of DNNs), field-programmable gate arrays (FPGAs) and other application specific integrated circuits (ASICs), as well as by novel computing technologies including quantum computing (which replaces the binary states in digital computers with superpositions of binary states), optical and photonic computing, magnetic computing, memristors [78–80], nano-oscillators [81, 82], Ising machines, in-memory computing (which would remove the Von Neumann’s bottleneck [75]), spintronics, and even more exotic approaches like DNA or thermodynamic [83, 84] computers. Some of the first unconventional accelerators are also already reaching the market. As an example, we can consider the dedicated integrated photonic circuit operating as an ultra-fast quantum random number generator [85, 86] developed by a Spanish spin-off startup called Quside.

In addition to the aforementioned approaches, there is also the field of neuromorphic engineering. Inspired by the extraordinary capabilities and power efficiency of the human brain and promising a tailored-made solution for AI computation, neuromorphic chips directly benefit from all the previously discussed factors: growth in AI as well as the shift towards application-specific accelerators and new computing paradigms. In conclusion, quoting again from Yann LeCun: *"We definitely need deep-learning focused hardware, especially since DNNs seem to perform better the larger they get."* [64] - and neuromorphics offer one of the most exciting answers to this call to action.

2.2 Neuromorphic engineering



Neuromorphic engineering is an interdisciplinary field of study that combines knowledge from neuroscience, computer science, and engineering to design, build and deploy neuro-inspired computing systems, with the ultimate goal of reaching the information processing abilities and efficiency of the human brain.

The case for building brain-inspired hardware was first made by Carver Mead from Caltech in his book "Analog VLSI and Neural Systems" [87] back in 1989, stating that *"The nervous system of even a very simple animal contains computing paradigms that are orders of magnitude more effective than are those found in systems made by humans."* Research of Mead's former students further laid the groundwork for the field. This includes Mischa Mahowald's work on silicon retina modules and Kwabena Boahen's work on Stanford's Neurogrid system. Nowadays, neuromorphic engineering is a multifaceted, rapidly growing engineering field, with an expected compound annual growth rate (CAGR) between 20% [88] and 45% [89] and with an expected global market of \$20 billion [88] by year 2030.

2.2.1 Definition

As a rapidly growing and constantly developing field with significant promise (and a futuristic name), neuromorphic engineering nowadays encompasses a very wide range of different approaches, technologies and systems. As such, properly assessing if a given technology is "neuromorphic" may sometimes constitute an interesting challenge [90].

Fundamentally, the term "neuromorphism" implies an isomorphism with biological neurons, particularly in terms of their behaviour and dynamics. This can be considered as a "bottom up" approach, where the starting point is a neuron and/or a synapse, either biological or engineered. From this starting point, the aim is to scale up towards a large scale system, such as a full physical neural network. Bespoke electronic circuits can be designed that mimic the dynamical behaviour of biological neurons to an extent given by the circuit design [91]. Some works have also proposed the name "biomorphic" [92] for such approach.

More loosely, neuromorphic hardware refers to systems that directly incorporate the fundamental building blocks of artificial neural networks (such as artificial neurons, axons or synapses) at the hardware level. Here, various degrees of abstraction from biology can be employed: some systems operate in an event-based fashion, some are mixed-signal, and some consist of a large number of small, simpler processors (so called manycore systems [93]). These are currently the most powerful and largest scale neuromorphic systems.

Finally, there are also approaches that can be considered the most on the top-down methodology, where the starting point can be an algorithm (such as matrix-vector multiplication that underpins modern deep learning). Based on the algorithm, a hardware is tailor-made to perform a particular operation at high speed and efficiently. Such devices are typically called AI accelerators, and in the author's opinion should be considered as a different (but not any less powerful) class of devices from neuromorphic chips. This conceptual difference is also shown in Fig. 2.3.

Given the broad range of different approaches to brain-inspired computing, these architectures are likely to provide solutions for different use cases with various system

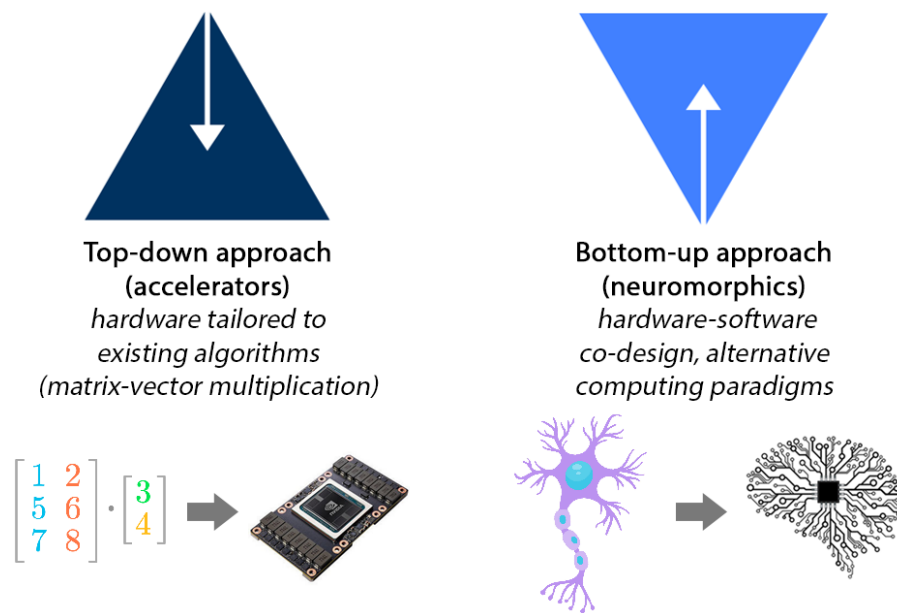


Figure 2.3: Comparison between top-down and bottom up approaches to designing AI-specific computing architectures.

and biological isomorphism requirements [10]. An excellent overview of neuromorphic approaches in hardware is available in [94].

2.2.2 Characteristics of neuromorphic systems

The main motivation for neuromorphics comes from the desire to build application-specific computing hardware for efficient realisation of AI and ML algorithms. As previously discussed, the goal is not to supersede digital computers, but rather augment them for the tasks where traditional digital von Neumann architecture provides sub-optimal performance. Edge-computing is also an area which is poised to particularly benefit from efficient, fully local AI computation, as well as the field of neuroscience, which can use dedicated, high-performance hardware to further our understanding of the brain [95], a task that is exceedingly challenging on digital computers.

Neuromorphic computing systems typically employ some common principles. These include a very high degree of computing parallelism, mixed digital-analog representation of signals, adaptive & local learning rules, co-location of memory and computing and event-based signalling. Furthermore, asynchronous operation and use of the continuous

temporal dimension (which is not typically used to represent data in digital computers besides multiplexing) are other features which are commonly explored in neuromorphic hardware. Unlike in the (exact) digital computers, a certain degree of randomness is inherently present in all biological brains. Some argue that this stochastic nature is a feature of its own, and is also represented in some neuromorphic designs [96]. Finally, there is also ongoing research focus on utilising and employing more unconventional neural network learning rules, including unsupervised [64] or local learning.

One of the important challenges for creating practical neuromorphic chips lies in efficient interfacing of these systems (which may operate with non-standard or analog represented data) to digital computers. Event-based computer vision systems, where data is represented in sparse events (spikes) continuously and with high temporal resolution, yield themselves as the most promising candidates for neuromorphic processing. The main advantage of these bio-inspired vision sensors lies in their sparsity, lower data volumes and their reduced or suppressed motion blur [97], among other benefits. These systems readily yield themselves suitable for combining with deep learning approaches [98], convolutional neural networks [99] and spiking neural networks [100] for data-efficient computer vision solutions.

The key motivation behind research papers on neuromorphic chips has been evaluated by [94], demonstrating how the trends have shifted towards low-energy computation. These trends in keyword usage are depicted in Fig. 2.4.

2.2.3 Neuromorphic electronic platforms

There is a wide range of both large and small commercial entities as well as universities and research institutes that are developing neuromorphic platforms in electronics. The two main directions of research are either towards datacenter scale chips/systems (optimisation of general purpose workloads) or embedded/edge computing approaches, which typically focus on lower-precision, and specific tasks/inference with high energy efficiency.

Neurogrid [101] was developed by Stanford University and is realised using a mixed-signal architecture with analogue circuits providing neuronal and synaptic dynamics,

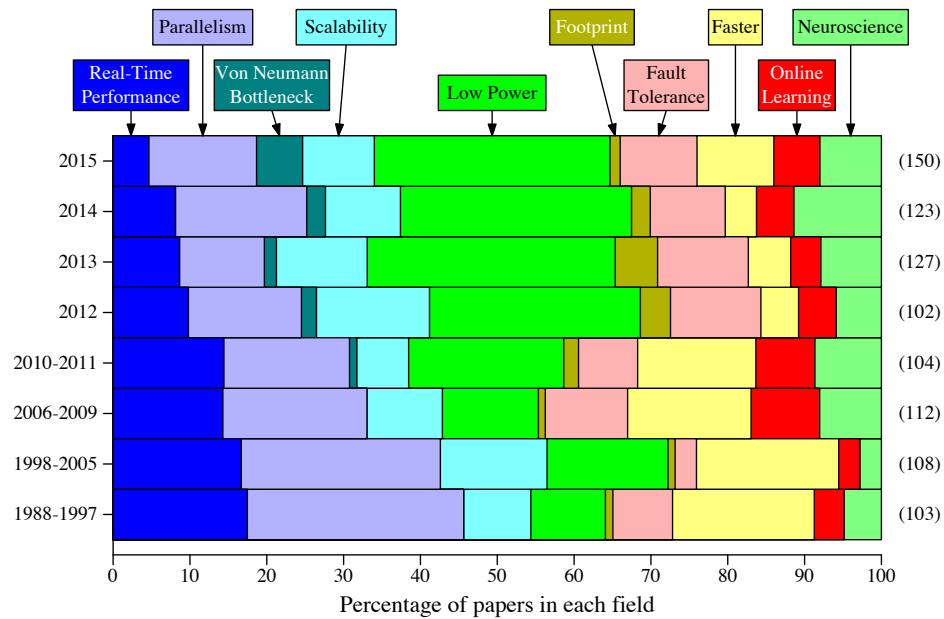


Figure 2.4: Keyword (research motivation) trends in publications over time. Reproduced from [94].

while utilising digital communication with spikes. SpiNNaker2 by the University of Manchester [102] represent a manycore system, based on a large number of low-power ARM architecture processors (10 million Cortex-M4F units [103]). It is therefore an all-digital approach that is fully programmable, as each processing element contains the processor, local static RAM (SRAM) and multiple hardware accelerators [104]. SpiNNaker systems are particularly useful for very large networks [105]. Finally, BrainScaleS-2 by the University of Heidelberg [106] is a mixed-signal system that features 512 analog, adaptive, fixed dynamics LIF neuron circuits and allows for operation at speeds more than $1000\times$ higher than biological timescales. Each neuron integrates stimuli from 256 6-bit synapses, and the system allows for low latency operation [106]. An edge-application focused processor operating over second-timescales with millisecond precision and online learning capability was recently presented in [107]. Among more recent entrants to the field is Tsinghua University and their Tianjic processor [108], whose unique selling point is its capability to combine in one architecture both spiking and conventional deep neural networks. The chips contain 156 neural cores, capable of

simulating 40,000 neurons and 10 million synapses.

Intel has recently introduced a second generation of their *Loihi* chip (Loihi 1 [93, 109], Loihi 2 [110]), which provides a very high degree of flexibility thanks to its fully programmable neuronal dynamics capability. In true neuromorphic fashion, it relies on event-based, sparse and time-coded signalling and offers up to one million neurons and 128 million synapses on one chip. A similar number of neurons and synapses is available in IBM's all-digital *TrueNorth* architecture [111] (4096 neurosynaptic cores with 256 fixed LIF neurons [112], with 256 low-precision synapses per neuron). Besides these major corporations, neuromorphic research efforts are also led by companies such as BrainChip with their event-based processor *Akida* [113] or SynSense (formerly aiCTX), a spin-off from the Institute of Neuroinformatics in Zurich, with their family of dynamic neuromorphic asynchronous processors (DYNAPs) [114] including the Dynap-CNN chip, marketed as efficient solution for CNNs converted to SNNs, and Speck, an event-based vision processor. Other market players include Axelera with a mixed signal, in-memory approach in their *DIANA* processor [115]. In-memory approaches in particular seem to be gathering growing attention, with recently announced NeuRRAM architecture [116] based on analog resistive RAM utilising 3 million memory cells and thousands of neurons. A comprehensive overview of the major large-scale neuromorphic electronic platforms (with comparisons) is provided in [117], with additional up-to-date details available in [118].

Finally, a brief focus should also be provided to the software side enabling the function of neuromorphic hardware and advances in neuroscience. *Lava* is a toolkit used to program Intel's Loihi processors, *sinabs* [119] or *Rockpool*, which can be deployed on SynSense chips, and *Nengo* [120], which can be used to program the SpiNNaker systems [121]. Another hardware-focused toolkit is *RANC* [122], inspired by TrueNorth architecture and focused on open-source implementation with FPGAs. Spike-based machine learning frameworks include *BindsNET* [123], *snnTorch* [124], *Norse* [125], *spikingJelly* [126] and *Tonic* [127]. An extended up-to-date list of spiking-related resources is maintained at [128].

2.3 From electrons to photons: Neuromorphic photonics

Optical computing attempts date way back prior to the current AI-fuelled neuromorphic photonic computing boom. Already in 1985, the prospects of optical neural networks were considered [129]. Soon after, researchers proposed programmable, optoelectronic multi-processor architectures focused on achieving high computational parallelism through augmentation of electronics with holographic, free-space optical interconnects [130] and multi-channel, dual-polarisation wavefronts [131]. A multi-layer optical neural network was shown in 1996 [132] with nonlinear activation functions achieved by liquid crystal valves. A few years later, an all-optical neuron design was proposed based on a laser diode with external optical feedback [133], demonstrating a nonlinear response (activation function) and both excitatory as well as inhibitory functionality. Furthermore, the dynamical isomorphism between excitable lasers and integrate-and-fire neuronal models was first reported in [134], making the case for realisation of photonic neurons even stronger. In the following years, many different approaches would be introduced for neuromorphic functionality. Among the first were lightwave systems, for example based on chain of electro-absorption modulators and nonlinear loop mirrors [135] for pulse integration and thresholding. This was followed by a similar study employing a more complex layout with five signal processing stages to realise all-optically a spiking LIF photonic neuron [136].

Nowadays, neuromorphic photonics is a booming research field, primarily fuelled by two main factors: i) astonishing developments in the domain of AI and ML leading to significant (and growing) energy requirements of large-scale AI models, b) advances in the field of photonic integration, which significantly increases the prospects of practical, light-powered computing [137]. For an overview of the fundamental principles of the field, see [138].

2.3.1 Advantages of photonics for neuromorphic systems

There is a multitude of properties typically exhibited by photonic and optical systems that are highly desirable for neuromorphic computation. These include:

- The use of photons for information transmission and signalling. Using these (typically) mutually non-interacting bosons solves some practical constraints of using electrons (such as direct interaction between signals) while simultaneously introducing new physical degrees of freedom for information multiplexing (i.e. wavelength, intensity, polarisation, angular momentum [139]) which can increase the communication bandwidth and be further utilised during computation.
- Optical signals propagated through waveguides exhibit less attenuation and generate significantly lower amount of heat (for off-chip light sources) [138], which may allow for tight component packing and three-dimensional layouts. In comparison to CMOS gates, where energy consumption is dominated by the need to drive gate-to-gate interconnects, optical guides carry signals passively and with lower latency and higher possible bandwidths. Optical signals propagate without requiring voltage, since light as an electromagnetic wave propagates without the need for a potential difference [140].
- Arguably the strongest case for photonics-based neuromorphic systems is their unmatched bandwidth. Optical systems can operate at bandwidths surpassing what is achievable in electronics, going beyond tens of gigahertz rates [141]. Due to its high bandwidth, optics allow us to use very short pulses, which are very difficult to propagate in large electronics systems [142].
- The propagation velocity of electrical signals through metallic interconnects scales depending on the capacitance per unit length: As a result, propagation speed is reduced for higher numbers of interconnections [143]. This is not the case for optical signal propagation, where velocity is independent of the number of interconnects.

- Optical systems may be capable of processing information passively solely by the physics of light-matter interaction [144]. For example, this can be seen in the interaction of light waves with inverse-designed diffractive elements [145].
- In the case of synchronous systems, the "clock rate" of an optical system is limited only by the laser modulation frequency, which can go above 100 GHz. Neuromorphic spike-based optical systems are typically asynchronous, avoiding operation speed reduction caused by the need for sampling and digitisation in clocked logic gates, while also reducing overhead energy consumption [146].
- Unlike digital neuromorphic systems, where neuronal dynamics are computed (simulated), photonic devices may naturally exhibit a full set of dynamical neuron-like properties (like is the case of injection-locked VCSELs, Chapter 3 of this thesis). This may renders analog circuits as more efficient for computing, while digital neuromorphic systems (with their precision and tuneability) may provide more suitable for e.g. neuroscience.

2.3.2 Disadvantages of photonics for neuromorphic systems

While photonics offers many highly desirable properties for implementation of neuromorphic architectures, certain challenges are also faced during development of such systems. These include:

- The additional degrees of freedom available by using photons provide additional parameters that need to be considered (or even controlled) during system design or operation.
- Variance between components. Particularly for very small components, individual variations due to their high sensitivity to changes in the fabrication process remain an ongoing practical challenge. This challenge is also present in analog electronics.
- Challenges in terms of fabrication, particularly for integrated approaches. However, the field of photonic integrated circuits (PICs) is developing rapidly, in part thanks to building on the foundations of mature CMOS (silicon) fabrication processes.

- For optoelectronics, the need for O-E conversion might place constraints in terms of bandwidth and speed [147]. For example, in the HolyLight system, photodetectors were found to be the energy consumption bottleneck, consuming more than 85% of the power [148].
- The dimension of photonic neurons and synapses is limited by the wavelength of the used optical signals ($\approx 1 \mu\text{m}$) [149]. However, there is research into sub- λ approaches that seeks to alleviate some of these limitations [150].
- As discussed by [151], the trainable parameters in optical modulator meshes act on the matrix elements in a highly nonlinear fashion, which may make the network training phase more challenging.
- Neuromorphic photonic systems are analog. Despite analog approaches being known and studied for decades, computing is clearly dominated by digital approaches. Arguably, this can be attributed to resilience of digital encoding to minor errors occurring during computation. In contrast, analog computing relies on continuous ranges of values rather than a discrete representation, and therefore can not rely on self-corrective error mechanisms, causing a loss of fidelity at every step. However, as spiking can be considered as a digital (in amplitude) encoding scheme, it might potentially be the key approach to overcoming this issue.



2.3.3 Taxonomy of light-powered neuromorphic approaches

The field of light-powered hardware designed for acceleration of computation of AI algorithms is vast, and sprawls across many different approaches, methodologies, devices and technologies. Therefore, providing a taxonomy for these approaches is rather challenging due to this high number of different methods for realising such hardware.

One option for broadly categorising is, as suggested by [152], to group the approaches in two main classes: a) used neuron type and b) neural network approach. An alternative way to classify different approaches is based on dividing them between all-optical and optoelectronic approaches. Optoelectronics benefits from combining both electrons and photons, each excelling at different key functionalities [153] with light providing the


capability to address bandwidth and efficient interconnectivity [142]. On the other hand, others claim that O-E conversion might place constraints in terms of bandwidth and speed [147]. General AI acceleration approaches based on optics can be classified as either being closer in function to conventional accelerators (typically continuous wave operation, focus on artificial and deep networks) or as closer to being purely "neuromorphic" (event-based operation in continuous time). A different perspective [154] proposes to group neuromorphic approaches based on being stateless (no memory) or stateful, and fully or locally interconnected. Finally, recent work [155] proposed focusing on devices in terms of their main functionality: either providing non-linearity (activation function), signal routing or synaptic functionality.



In the following overview, the abbreviation "ONN" stands for an "optical neural network", denoting an artificial neural network architecture that is realised in hardware. In the scope of this chapter, the term ONN is mainly conferred on continuous wave (non-spiking) approaches. Bibliographic sources related to non-spiking (continuous wave) approaches are highlighted using the following red symbol . Where spikes are underpinning the functioning of the device and/or network, the referenced publication will be highlighted with the following blue symbol .

2.4 Photonic neural networks: Integrated approaches

2.4.1 Mach-Zehnder interferometers (MZIs)

The general principle of operation of Mach-Zehnder interferometer (MZI)-based optical neural networks is well described, for example in [156]. In a hallmark manuscript that laid the foundation for the start-up company Lightelligence Inc., a feed-forward, fully connected neural network is demonstrated  [141], based on a 4×4 weight matrix in an optical interference unit (OIU) based on 56 programmable MZIs in the Reck architecture [157], and optical nonlinearity units (ONU) simulated in silico, modelling the behaviour of a realistic saturable absorber. The network achieved 76.7% vowel speech recognition

accuracy (against 91.7% in computer at the time), with two orders of magnitude faster processing and three orders of magnitude lower energy requirement. The PIC containing the photonic mesh is shown in Fig. 2.5. A single network layer building block based

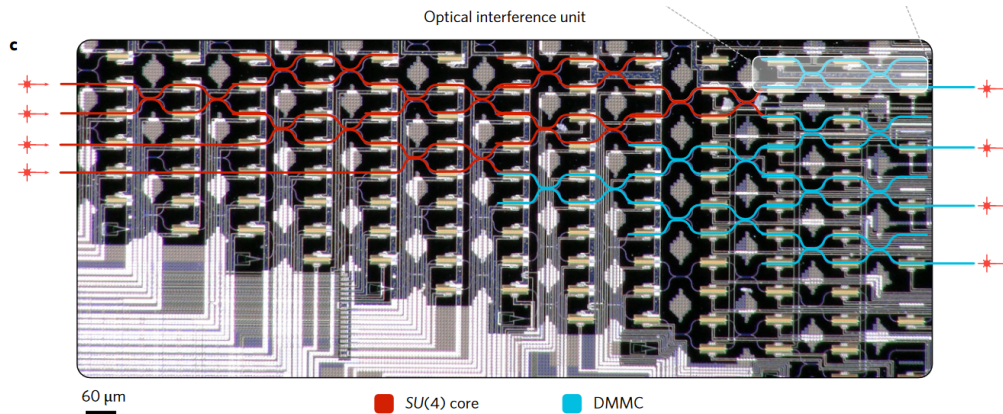


Figure 2.5: Illustration of the optical interference unit, which realises all-optically both the functions of matrix multiplication (highlighted in red) and attenuation (highlighted in blue). Reproduced from [141].

upon reconfigurable mesh of MZIs was also shown in [\[158\]](#). A similar approach incorporating back-propagation was demonstrated in [\[159\]](#), computing gradient terms by propagating an adjoint field and interfering its time-reversed copy with the original field. In this approach, a forward pass is performed and the optical signal is measured, then the error signal is sent back and the adjoint is measured. Finally, the sum signal is forward passed and its gradient is computed via digital subtraction. Rigorous analysis of the in-situ backpropagation method is also presented in [\[160\]](#), demonstrating how offloading the computationally demanding matrix-vector products and matrix gradient computations (backward pass) provides a computational advantage. Besides backpropagation, a mesh (feed-forward network) of MZIs can also be trained using a two-step ex-situ procedure combining the stochastic gradient descent (SGD) algorithm with a genetic algorithm [\[161\]](#). In [\[162\]](#), an optoelectronic approach is used, with a mesh of integrated MZIs connected to nonlinear activation function blocks (neurons). These neuronal blocks provide a nonlinear response by splitting the optical input signal, converting a small portion of it into an electrical signal, and then using it to modulate the remaining portion of the source optical signal in an another MZM. They numerically

(and later also experimentally [◆](#) [163]) demonstrated the nonlinear activation function of such arrangement and showed an optical neural network performing image (digit) recognition with an accuracy of up to 93.89% reported for a three-layer network model. A LeNet-5 (five layers) CNN was implemented in a MZI mesh, with kernel operations implemented in the optical domain, and nonlinearity and interconnectivity implemented in FPGAs [◆](#) [164]. Compared to previously mentioned Shen's work [◆](#) [141], the singular value decomposition (SVD) algorithm is replaced here with pseudo-real-value matrix decomposition, which decreases the mesh size (and therefore the influence of fabrication imperfections).

In the MZI-based integrated architectures, thermal cross-talk between modulators represent a common challenge lowering the performance [141, 158], as do fabrication process variations between devices and manufacturing defects or absorption and coupling losses [165]. Some works have proposed self-correcting mechanisms [166] to mitigate the unwanted influence coming from these effects.

In addition, a simpler experimental layout, utilising WDM signals and TDM multiplexing of input data into a single MZM [◆](#) [167] relies on dispersive shifting of individually modulated channels and their readout on a photodetector. This MZM-based perceptron was used with 49 WDM channels to process MNIST handwritten digit recognition and perform cancer-cell recognition, both with good degree of accuracy.

2.4.2 Semiconductor optical amplifiers (SOAs)

An integrated, all-optical neuron based on a semiconductor optical amplifier (SOA) has been demonstrated to exhibit a sigmoid-shaped activation function [◆](#) [168]. Using a differentially-biased SOA-MZI followed by a SOA operated in small-signal regime as wavelength converter, the logistic sigmoid transfer characteristic has been validated using four WDM inputs (1550 nm). In [◆](#) [169], an array of four-to-one weighted addition modules has been experimentally demonstrated using SOAs integrated on an InP platform. By calibrating for the nonlinearity of the weight-SOA gain curve, they were capable of utilising the SOAs in a three-layer network (with electronically implemented nonlinearity) to successfully perform the Iris classification task [◆](#) [169]. The experimental

setup is shown in Fig. 2.6 Furthermore, utilising the effects of nonlinear polarisation

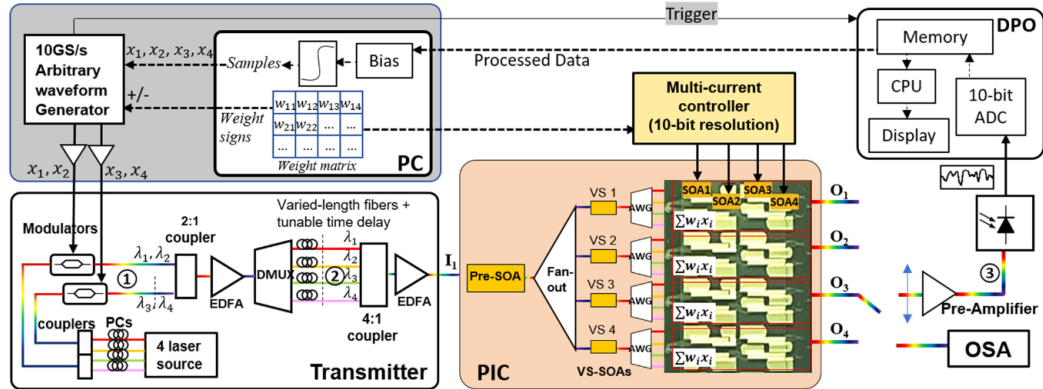


Figure 2.6: Experimental setup of SOA-based optical neural network. The Transmitter block shows how the four signals are modulated and multiplexed prior to entering the PIC. PCs: polarisation controllers; DMUX: De-multiplexer; Pre-SOA: Semiconductor optical preamplifier; VS: Input vector selector gate; AWG: Array waveguide grating; EDFA: Erbium doped fibre amplifier; PD: Photodetector; OSA: Optical spectrum analyzer; DPO: Oscilloscope (OSC). Tunable time delay indicated with dashed lines. Reproduced from [169].

rotation and cross-gain modulation within an individual SOA, both excitatory and inhibitory aspects of a spike-timing dependent plasticity (STDP) functionality have been demonstrated [170]. In comparison to other integrated photonic devices, SOAs have a fairly large spatial footprint (with device length in mm range), which may pose as challenge for architectures with high required integration density.

2.4.3 Microring resonators (MRRs)

Alongside optical modulators, microring resonators (MRRs) are arguably the most studied integrated components for optical neural network architectures, often operated with WDM optical signals thanks to their controllable frequency selectivity.

Continuous wave & ONN approaches

In a top-down driven fashion, the multiply-and-accumulate (MAC) algorithm (which underpins modern deep learning) was demonstrated in an integrated optical signal processor utilising an array of 4×4 MRRs and achieving 80 megaMAC/s at 10 MHz

clock rate \blacklozenge [171]. In \blacklozenge [172], a pair of balanced photodetectors receives WDM optical inputs, and provides the summed signal via an electrical wire to a microring-based, modulator-class neuron which re-modulates the summed signal onto a new wavelength. The MRR contains an embedded p - n modulator and n -doped heater realised in silicon photonics. Among other functionalities, coincidence detection of ≈ 5 ns pulses and autaptic (self)cascadability are demonstrated. The microrings also yield themselves to operation as weight banks $\textcircled{\bullet}$ \blacklozenge [173–175], providing the key synaptic weighting functionality (particularly for WDM-based approaches) on integrated platforms that can scale to larger networked arrangements. Finally, a complex optoelectronic accelerator named *HolyLight* has been demonstrated in \blacklozenge [148], utilising an architecture based on photonic adders, shifters and matrix-vector multiplier units. These units are based on add-drop microdisks, and the architecture is designed to efficiently compute CNNs.

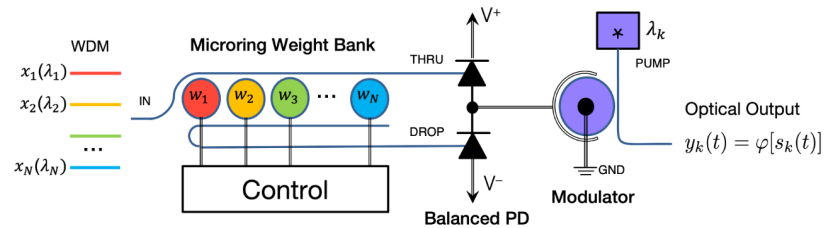





Figure 2.7: Silicon photonic non-spiking neuron implementation using a microring resonator weight bank. Reproduced from [176].

Spiking approaches

Besides continuous wave approaches, excitability and spiking has also been demonstrated in microrings. In $\textcircled{\bullet}$ [177], both experimental and numerical evaluation was performed for a semiconductor ring laser (SRL), showing single- and multi-pulse excitability without the need for feedback mechanisms, but with a high degree of variation in the produced spikes. A SRL under optical injection was also shown in theory to exhibit controllable excitability $\textcircled{\bullet}$ [178]. Passive silicon-on-insulator microrings under external light injection can also exhibit self-pulsing and class II excitability thanks to nonlinearities, as demonstrated via modelling in $\textcircled{\bullet}$ [179, 180]. A later study has validated that such excitable SRL exhibits all hallmarks of neuron-like excitability, including thresholding, refractoriness,




cascadability, leaky integrating dynamics and inhibitory spiking behavior  [181]. More recently, excitable behaviour was shown in coupled pair of individually detuned Kerr resonators with a single laser beam used as a driving signal  [182].

Thanks to the capability of microrings to behave both as spiking and weighting elements, an all-microring based spiking photonic architecture can be realised. A theoretical model of such system has been reported in  [183], demonstrating, in addition, STDP functionality and successfully implementing a spike sequence learning task via a remotely supervised learning approach.

In conclusion, microrings represent a promising candidate for integrated neuromorphic photonics due to various feasible use cases and their favourable spatial footprint. Similarly to other integrated photonic components, the high sensitivity of microring resonators to crosstalk, as well as inconsistencies and insufficient process control during fabrication remain a challenge for device stability.

2.4.4 Electro-absorption modulators (EAMs)

Electro-absorption modulators are devices that absorb light depending on the applied voltage, as defined by their non-linear absorption curve. They are based on the Franz–Keldysh Effect (a change in the absorption spectrum by modulating the bandgap energy using external bias) and are likely to have lower spatial footprint than interferometer-based modulators, enabling higher packaging density [184].

A WDM-based, electro-optic neuron based on four building blocks (photodiode, transimpedance amplifier, electro-optic modulator, CW laser) has been theoretically reported in  [184]. This work provides good details on coupling methods between the optoelectronic elements and demonstrates a computer-based deep neural network performing the MNIST hand-written digit image classification task using the nonlinear activation function exhibited by the system. The architecture from  [184] is shown in Fig. 2.8. A leaky ReLU optical neuron is realised using an ITO-graphene based heterojunction electro-absorption modulator  [185]. Furthermore, this nonlinear neuronal unit is used in a WDM-based broadcast-and-weight networked scheme utilising ring modulator weights. To validate the prospects these devices in an ANN setting,

this achieved transfer function was implemented in a three-layer ANN model (using the Keras ML framework) and demonstrated classification of digits from the MNIST dataset [185]. In more recent work [186], high bandwidth SiGe EAMs were demonstrated to

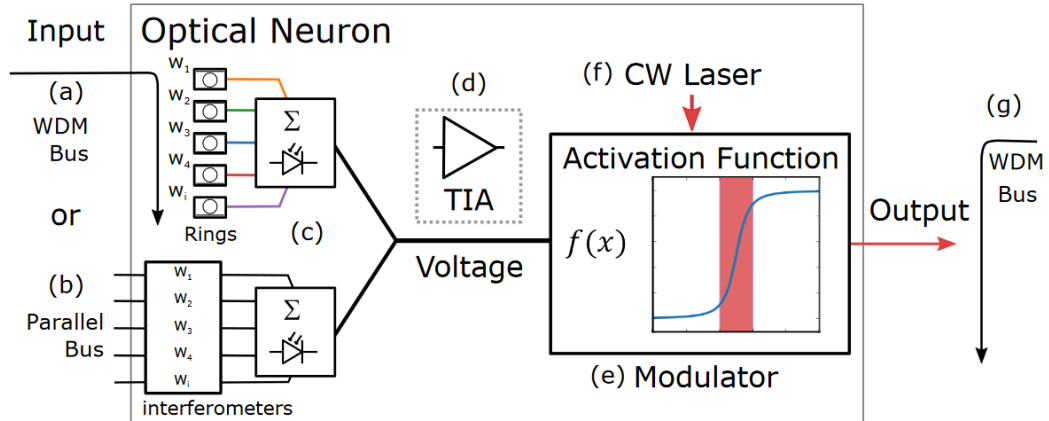


Figure 2.8: An electro-optic neuron layout. Input is provided by (a) an array of rings or (b) with an interferometer network. The neuron sums the optical signal with a photodiode converting the signal to a voltage (c), optionally amplified by a TIA (d), drives an electro-optic modulator (e) modulating a CW laser (f), which produces a nonlinear transfer function at the output (g) of an photonic neural network (NN). Reproduced from [184].

be capable of performing both input data and network weight updates at 16 GHz. They perform the MNIST dataset digit recognition using a 5-layer neural network, where the last two cascaded neural network layers are realised in photonic hardware. Using this system, the achieved MNIST digit recognition accuracy was 99.18%.

2.4.5 Phase-change materials (PCMs)

An integrate-and-fire neuron was proposed in [187], consisting of an integration unit (two double bus ring resonators with a germanium-antimony-tellurium (GST) phase-change material (PCM) cell) and a firing unit, consisting of a photonic amplifier, a circulator and a PCM-integrated rectangular waveguide. This work studied in theory the operation of such a node for rate-coded input spike trains, and validated the integrate-and-fire behaviour.

An all-optical neuromorphic chip was demonstrated by [188], using 4 neurons

and 60 synapses based on phase-change materials (PCM). This systems was capable of both supervised and unsupervised pattern recognition. The *integrate-and-fire* optical neurons were realised by combining micro-ring resonators with optically modulated PCM cells. To enable unsupervised learning of the network, a feedback mechanism was implemented by sending the output signal back into the PCM-based attenuators on the waveguides (photonic synapses).

In follow-up work, a WDM-based photonic tensor core \blacklozenge [189] utilising PCMs as non-volatile network state memory elements has been proposed. This enabled an architecture capable of performing a convolution functionality at very high throughput and processing speed with low power consumption. In this work, the nonlinearity was provided after the convolution steps in a computer. A scheme describing the architecture is shown in Fig. 2.9. The PCM-based integrated platform was further

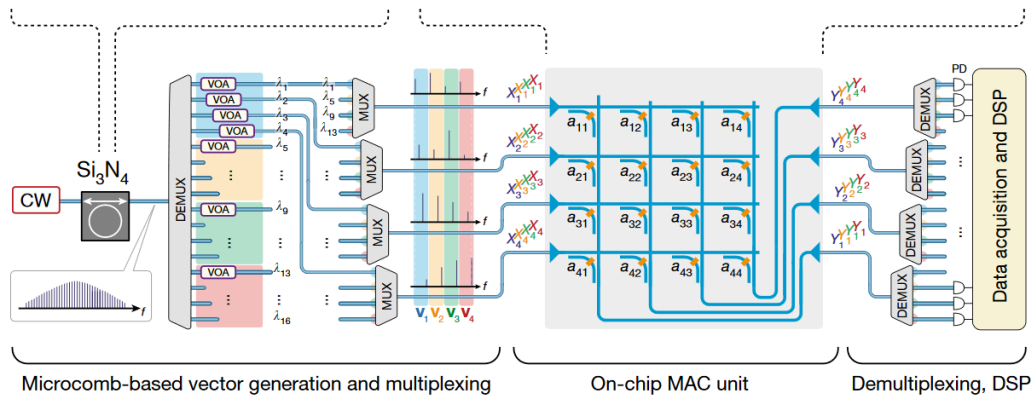


Figure 2.9: Scheme of the multiplexed all-optical PCM-enabled neural network (MAC unit). The input vectors are using a comb source. The entries of different input vectors are grouped together again employing wavelength multiplexing and sent to the on-chip MAC unit that performs the calculations. After combining the correct wavelengths with optical wavelength division demultiplexers, the multiplication results are obtained from the photodetectors followed by digital signal processing. Reproduced from [189].



utilised for demonstration of a specific supervised learning algorithm based on association between two independent inputs \blacklozenge [190]. As more pairs of pulses (input and teacher signals) arrive at the associative element, the PCM cell gradually switches from the crystalline to amorphous state, strengthening the weight between the two inputs. Using a network of four associatively-trainable PCM units, the network is trained to recognize

images of a particular kind (cat images). It is shown that the system is capable to classify cat images from dog images.

The use of PCM enables long-term non-volatile storage of synaptic weights, but simultaneously raises some challenges, such as the requirement of high input power precision to successfully operate the PCM units, operation rate limits due to crystallization speed of neuronal PCM cells, and possibly limited lifespan due to repeated material melting after every excitatory event [191].

2.4.6 Josephson junction-based optoelectronics

Josephson junctions are made by sandwiching a thin, non-superconducting material layer between two superconducting materials. Interestingly, they share some history with resonant tunnelling diodes, one of the main devices studied in Chapters 4 and 5 of this thesis.[†]

Optoelectronic approaches based on superconducting nanowire single-photon detectors  [193] offer some very desirable properties, including very high photosensitivity, zero static power dissipation and very high energy efficiencies reaching 20 aJ/synapse event. In this theoretical work, a system based on such technology is thoroughly analyzed, including its integrate-and-fire functionality and few-photon spike activation capability, as well as various neural node and network architectures. Inevitably, a superconducting approach requires ultra-low temperature cooling (here at approx. 2 K). Nonetheless, even the power expense of cooling the system down might prove tolerable given the system's extraordinarily low power requirements, assuming a good degree of scaling is achieved. Further details on this class of optoelectronic neuromorphic systems are provided in a recently published review article by the team at the National Institute of Science and Technology (NIST) in the US  [194].

[†]The 1973 Nobel Physics Prize for discoveries of tunnelling phenomena in solids to Drs. Leo Esaki, Ivar Giaever and Brian Josephson [192]. Interestingly, Esaki's diode can be considered as a predecessor to double barrier quantum well resonant tunnelling diodes (RTDs).

2.5 Photonic neural networks: Spatial approaches

Optical wave propagation and diffraction through a metamaterial (an array of rectangle slots) was shown in theory to passively perform matrix-vector multiplication in [\[195\]](#), with low energy requirements. Furthermore, convolution-like functionality was numerically demonstrated in [\[196\]](#) using a single metasurface (a layer of 100 nm silicon nanodisks), enabling an all-optical edge detection functionality in images. Signal processing analogous to neural network computing has been modelled for wavefronts passing through nanophotonic media [\[197\]](#) and metasurfaces [\[198\]](#). Using sub- λ scatterers, it allows for ultrahigh computing density within a medium as small as tens of wavelengths. Experimentally, a diffractive processing unit (DPU) was developed in [\[199\]](#), where almost all required operations are implemented with optical diffraction. The DPU also supports operation with millions of neurons. Multiple architectures are proposed, including deep neural networks and networks-in-networks. The input data are converted into complex-valued input fields electro-optically using digital micro-mirror devices (DMDs) and spatial light modulators (SLMs). They have demonstrated processing of the MNIST digit classification task with 940 thousand neurons, and achieved 97.6% inference accuracy with adaptive training to avoid error accumulation between layers. A scheme and visualisation of a deep diffractive processing unit from [\[200\]](#) is shown in Fig. 2.10.

2.6 Photonic neural networks: Laser-based approaches

In this section, optical approaches based on excitability and neuron-like spiking will be discussed, with a main focus on these phenomena in lasers.

A first option is to employ optoelectronic systems, such as a three-transistor LIF-mimicking circuit with dual PD inputs (excitatory + inhibitory) and a single coupled laser as optical output [\[152\]](#), operating experimentally at ≈ 10 kHz firing rate. This work also introduces a prospective, nanoscale integrated 10 fJ/bit level implementation of such a system with photonic crystal-based photodetectors and quantum-dot (QD)

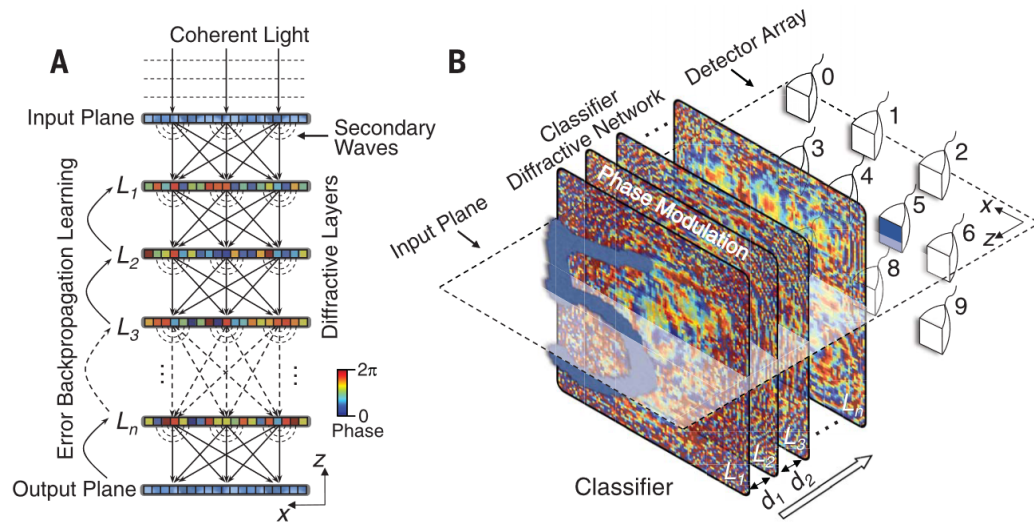


Figure 2.10: (A) A deep diffractive processing unit is based on multiple transmissive (or reflective) layers, where each point on a given layer acts as a neuron, with a complex-valued transmission (or reflection) coefficient. The transmission or reflection coefficients of each layer can be trained by using deep learning to perform a function between the input and output planes of the network. After learning, the layers are 3D-printed and perform a specific task at the speed of light. (B) visualisation of a classifier for handwritten digits. Reproduced from [200].

lasers. Chapters 4 and 5 also focus on neuromorphic functionality in optoelectronic circuits, realised with an excitable element (resonant tunnelling diode) driving a laser.

Regarding all-optical approaches, excitability has been demonstrated for example in QD lasers. A theoretical study demonstrating integrate-and-fire functionality in InAs/GaAs optically injected QD laser is provided in [201], with demonstration of ≈ 100 ps optical spikes. Experimentally, mode-locked QD lasers [202] were successfully utilised to demonstrate excitability and continuously firing, ≈ 70 ps dynamical regimes by means of amplitude modulated optical injection. Finally, an overview of excitability in semiconductor lasers with optical injection (in particular VCSELs) is provided in Sec. 3.2.

Chapter 3

Photonic neurons based on VCSELs: characterisation and applications

Relevant authored peer-reviewed manuscripts:

- ◆ Hejda, M., Robertson, J., Bueno, J. & Hurtado, A. *Spike-based information encoding in vertical cavity surface emitting lasers for neuromorphic photonic systems*. *Journal of Physics: Photonics* 2, 044001. (Invited). [10.1088/2515-7647/aba670](https://doi.org/10.1088/2515-7647/aba670) (2020)
- ◆ Hejda, M., Robertson, J., Bueno, J., Alanis, J. A. & Hurtado, A. *Neuromorphic encoding of image pixel data into rate-coded optical spike trains with a photonic VCSEL-neuron*. *APL Photonics* 6, 060802. (Invited). [10.1515/10.1063/5.0048674](https://doi.org/10.1515/10.1063/5.0048674) (2021)

THIS chapter introduces a class of semiconductor lasers called vertical cavity surface emitting lasers (VCSELs). Under a specific set of operational conditions, these lasers can operate as high-speed, all-optical spiking excitable devices, which makes them very promising candidates for artificial photonic neurons. This chapter provides experimental results with VCSELs, demonstrating their characteristics as optical spiking sources and their functionalities for spike-based neuromorphic information processing.

3.1 Fundamentals of VCSELs

In this thesis, most experiments have been conducted with a particular type of semiconductor laser called *vertical cavity surface emitting laser*, or VCSEL in short. As the name suggests, VCSELs emit light normal to the surface of the chip they are fabricated on. This imposes additional complexity on the epitaxial growth process due to requirements for precise control of layer thickness and quantum well alignment. Despite this, their structure and vertical emitting nature brings some significant advantages:

- VCSELs can be tested directly at wafer level without dicing and cleaving thanks to their facets already being cleaved. This makes their processing subsequent to growth considerably simpler, and significantly reduces their fabrication costs,
- they can also be manufactured monolithically in large arrays and ensembles, further reducing the costs,
- they have low threshold for lasing operation and low power consumption,
- the properties of the produced beams (low divergence and circular beam) make them well suited for coupling into optical fibres,
- a very short cavity guarantees single longitudinal mode operation [203],
- the low gain medium volume yields VCSELs for dynamic operation [46], allowing for modulation at high speeds.

3.1.1 VCSEL characteristics

The epitaxial growth for VCSELs is typically realised using molecular-beam epitaxy (MBE) or metal-organic chemical vapour deposition (MOCVD), and often relies on quantum well regions for enhanced performance. Due to their vertical emitting character, the length of the optical gain region is approximately equivalent to the active layer thickness (in the order of tens of nanometers to few μm [46]), shorter than in other types of laser diodes. Therefore, to achieve low threshold operation (in the mA range), very

Table 3.1: Advantages of VCSELs compared to EELs (left) and communication-type LEDs (right), Reproduced from [204].

VCSELs vs. edge-emitting lasers	VCSELs vs. LEDs
Low threshold current	High modulation bandwidth
High efficiency @ low powers	Focused output beam
Slowly diverging circular beam	Coherent output, narrow spectrum
Low-cost, wafer-level testing possible	Low operating currents
Simplified packaging, mounting	High output powers
Allow for 2D array operation	High power conversion efficiency

high reflectivity (99.5%) is required for the resonator mirrors to balance the short axial length of the gain region. This is achieved by multi-layered reflectors called distributed Bragg reflectors (DBRs), which are based on a stack of alternating epitaxial layers with different refractive indexes. Table 3.1 summarises the main advantages of VCSELs when compared to other semiconductor lasers (edge-emitting lasers) and light emitting diodes (LEDs) as outlined in [204].

The fundamental mode of a VCSEL is typically a highly linearly polarised TEM mode along one of two perpendicular directions, with a second higher order mode lasing with polarisation direction orthogonal to that of the fundamental [205]. Due to cavity anisotropies, the resonant frequencies (due to material birefringence) and amplitudes (due to dichroism) of these two modes exhibit slight differences [206]. Furthermore, thanks to the low thickness (typically around one wavelength) and emission along the z -direction, the difference between *up*- and *down*- spin states during carrier recombination must be taken into account [46]. When spin-up or spin-down recombination occurs [206], light with either left- or right-circular polarisation respectively is produced. These spin-based phenomena result in a rich variety of polarisation-related dynamics in VCSELs [207].

There are various III-V material systems used for VCSELs. These include GaAs/Al-GaAs ($\lambda \approx 850$ nm), and longer-wavelength emitting systems in InP, InAs or dilute nitrides, suitable for longer-distance, single-mode fibre-optic communications in the optical O-band (≈ 1310 nm) and the optical C-band (≈ 1550 nm). Such longer-wavelength lasers can nowadays be commercially sourced from telecommunications-sector manufac-

turers such as RayCan [203], Bandwidth10, Alight Technologies (technology acquired from Infineon), RPMC Lasers, Vertilas GmbH or Lumentum. Modulation rates and bandwidth are key for operation as telecom optical transmitters, and bandwidths reaching 30 GHz have been demonstrated in VCSELs [208] with 80 Gbit/s non return-to-zero data transmission rates [209].

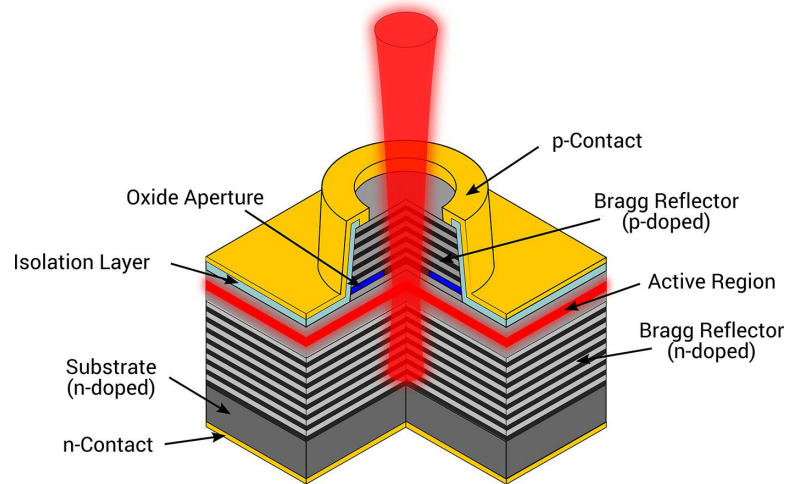


Figure 3.1: A schematic of a typical VCSEL device structure, consisting of an active semiconductor medium between two DBRs. Reproduced from [210].

3.1.2 Injection locking of VCSELs

If an oscillator is directly interfaced with another oscillator, they will mutually interact. In particular, if the two oscillators' frequencies are close, synchronization between the two oscillators may occur. This forms the foundation of the *injection locking* principle, where a slave oscillator is exposed to an external signal close to the natural, free-running first harmonic frequency of the slave, to alter the character of the oscillator's dynamics [211] by shifting and locking its oscillation frequency and phase to the master external signal. In lasers, injection locking can be utilised to stabilise single mode operation and achieve better performance, for example in terms of spectral narrowing, suppression of laser noise [212] and improving the modulation bandwidth of the device [213], making this approach highly interesting for the field of optical telecommunications.

To achieve injection locking, the detuning between the slave and the master signal

$\Delta\omega$ must be sufficiently small. The master signal must also be sufficiently strong. The detuning range around the cavity resonance is typically asymmetrical thanks to non-zero α factor introducing a red shift to the cavity mode [47]. In the parameter space [214], the slave laser remains locked to the master when operated at (or close to) stable solutions, with the boundaries between stable/unstable regions represented by bifurcations, and with nonlinear dynamics such as beat phenomena, oscillations, chaos [215] or excitability [216] possibly arising outside of the locking region.

Injection locking can also be utilised in VCSELs [203]. Due to the complex nature of light polarisation in this class of lasers, polarisation also needs to be taken into account when investigating injection in these devices. Since these lasers typically emit in two linear, mutually orthogonal polarised modes (a dominant and a subsidiary) that are slightly frequency shifted, injection can be performed into either of these two polarisation modes. At first, injection locking experiment in VCSELs were performed with injection into the subsidiary (orthogonally polarised) mode, and optically-induced polarisation switching was readily observed [217] with locking achieved in the subsidiary mode and suppression of the previously dominant mode. Furthermore, bistability was observed during injection locking for both the subsidiary [218] and dominant mode [219]. Parametric mapping for a C-band operating VCSEL was performed in [220] for both parallel and orthogonally polarised injection, demonstrating various kinds of nonlinear dynamics, including period one, period doubling and chaos. In these nonlinear dynamics via injection locking of one of the polarisations, the total output signal is often produced by the interplay of the two polarisation modes [221], with various degrees of interplay also governed by the injection detuning sign. Overall, these results demonstrate that VCSELs under optical injection offer a system with a rich and diverse range of nonlinear dynamics.

3.2 Spiking in injection-locked VCSEL neurons

Excitable (optical) spiking (see Sec. 1.2.3) represents a type of nonlinear dynamical behaviour that underpins the use of VCSELs as all-optical, artificial photonic spiking

neurons. In semiconductor lasers, excitability can be classified as either based on a saddle-node on invariant cycle (SNIC) bifurcation [222], or as excitability in the vicinity of a homoclinic bifurcation, which has been observed in devices with saturable absorbers [223].

Arising from the rich body of research focused on injection-locking in VCSELs and the corresponding nonlinear dynamics that can be achieved in these devices, neuron-like spiking has been reported as early as 2010 [224], using dual optical injection into a 1550 nm VCSEL. In a follow-up manuscript [225], injection locking of a linearly polarised signal into the VCSEL's subsidiary polarised mode was utilised to achieve polarisation switching and demonstrate a range of neuron-like spiking phenomena, including phasic and rebound spiking by the means of intensity modulation. This was further analysed theoretically in [226], revealing both single- and multipulse optical excitable responses as well as bursting dynamics under polarised optical injection in VCSELs. Later, a more detailed study [227] investigated and demonstrated precise, highly consistent triggering of sub-ns (≈ 100 ps) optical spikes in a 1310 nm VCSEL subject to both types of modulated, polarised optical injection (parallel and orthogonal). Furthermore, this work demonstrated the existence of a clear excitability threshold for (optical) injection perturbations. The findings on excitable spike activation in long-wavelength VCSELs reported in [227] provide the main foundation for the experimental techniques used in this chapter.

In an injection-locked VCSEL, the saddle-node bifurcation that gives rise to excitable behaviour exists at the boundary between the injection-locked and unlocked regions [43]. Additional approaches beyond intensity modulation of the injection signal have also been demonstrated for eliciting excitable responses in VCSELs. For example, in [228], a 980 nm VCSEL was subject to optical injection locking with controllable, 100 ps long phase-encoded perturbations. Similarly to the case of intensity modulation, thresholding for spike firing was also observed for the case of injecting phase perturbations. In a follow-up report, the existence of a refractory period for phase-modulated optical excitability was also confirmed [229]. Furthermore, excitability of both integrator and resonator types has been demonstrated in externally-injected VCSEL subject to phase

perturbations, including multipulse spiking, for different parameter settings [43]. More recent work has also demonstrated multi-frequency switching as a viable mechanism for triggering of optical spikes in an injection-locked 1550 nm VCSEL [230].



3.2.1 State-of-the-art in VCSEL-neurons

Controllable optical spike cascability and propagation was demonstrated in a feed-forward interconnected setup of two VCSEL-neurons [231], including propagation of inhibitory responses [232]. These functionalities were used to demonstrate an optical, toy-model analogue of retinal neuronal circuitry [233]. Spike propagation was studied in theory using the spin-flip model [234]. Besides modulation of injection signals, high-speed neuron-like excitable dynamics were also demonstrated for injection-locked VCSEL subject to electrical (bias current) modulation [235].

In recent years, advanced information processing functionalities have been investigated in optical spiking VCSEL-neurons. In [236], the leaky integrate-and-fire (LIF) functionality was experimentally demonstrated in a 1310 nm optically-injected VCSEL neuron. Based on this property, a first experimental proof-of-concept of high-speed, 4-bit time division multiplexed (TDM) pattern recognition task was realised (at 80 ps/bit rates), showing both repeatability and very high detection accuracy. In that work [236], the 4-bit digital patterns were pre-weighted (using a computer) prior to being input to the VCSEL-neuron. In [237], sequential convolutional functionality was demonstrated utilising a similar, TDM-based operational principle and using a second Mach-Zehnder modulator (MZM) to perform the operation in an all-optical fashion. Unlike in a typical convolution operation, where the value of the submatrix and kernel are multiplied, here the performed operation is realised using addition via passing the signal through two serially connected MZMs. The functionality was utilised for feature extraction in complex digital images, and demonstrated a very good degree of performance for edge-feature detection at high input data rates. Digitally pre-weighted convolutional functionality in the spiking VCSEL-neuron was also recently demonstrated in [238] for edge-feature extraction in digital images with very high reliability and at very high speeds (1.25 ns/pixel). This principle was further extended by combining the VCSEL-

neuron with a software-implemented spiking neural network (SNN) in [239]. In this work, the photonic VCSEL-neuron was used to perform high-speed edge-feature detection on a set of images of handwritten digits (MNIST handwritten digits dataset), and the output was fed to the software-implemented SNN to perform high accuracy digit classification. This demonstrated the prospects of a hybrid architecture, combining together physical (photonic) and software-based neuromorphic spike-processing techniques.

More recently, a single long-wavelength VCSEL was demonstrated as suitable for delivering a full photonic SNN, operating effectively as a spike-based analogue of an extreme learning machine (ELM) [240]. This has demonstrated a powerful optical data processing capability in an experimental layout with low overall complexity and hardware-friendly implementation. By combining the internal short-term dynamical memory of the spiking dynamical regimes in the VCSEL with TDM data processing approaches, optical spiking (binary) fingerprint traces can be obtained from the VCSEL, which allows for accurate classification of highly complex multi-dimensional data. In an alternative approach that extends beyond a single device and relies on an array of 5×5 VCSELs under injection locking [241], computation of deep neural networks was reported using space-time multiplexing and fan-out data copying. A single laser (in the centre of the array) propagates information with large fan-out while the remaining lasers encode information in their phase, and the resulting signal is collected on a photodetector array. For a more detailed overview of photonic neuromorphic and reservoir computing developments in VCSELs see [242].

Furthermore, besides the optical injection-locked approach, two-section VCSELs also have a strong demonstrated track record for use in neuromorphic functionalities. Such devices consist of two independent sections: a gain section (serving as a temporal integrator) and a saturable absorber (serving as a threshold detector), with the approximation of constant light intensity profile across the cavity. Self-pulsating in two-section devices can be described by the Yamada model  [243], which exhibits class I type excitability due to the presence of a saddle-node bifurcation  [244] (see Sec. 1.2.3). These devices have been demonstrated to exhibit analogous dynamics to leaky integrate-and-fire (LIF) neurons [245], and were further used in coupled, multi-device

arrangements to achieve spike synchronisation [246]. In terms of information processing in VCSEL-saturable absorber (VCSEL-SA) devices, supervised spike sequence learning was demonstrated in [247] and supervised pattern recognition based on spike timing (STDP) in [248]. In addition, a winner-takes-all spiking algorithm was reported in [249] and a two-layer spiking neural network was used to perform motion detection in [250] and sound azimuth localisation in [251]. Except for [245], all of the works mentioned in this paragraph have been realised in theory, with experimental reports still outstanding.

3.2.2 VCSELs used in this thesis

All the VCSELs used in this work were commercially sourced, single transverse-mode devices manufactured by the company RayCan. These were coupled with single-mode fibre pigtailed of either FC/PC (physical contact) or FC/APC (angled physical contact) connectors and packaged using the TO-56/TO-90 standard. VCSELs at both the O- (1310 nm) and C- (1550 nm) telecommunications bands were used in this work. The VCSELs' structure contains an InGaAs QW active layer together with an InAlGaAs/InAlAs distributed Bragg reflector (DBR). Typically, these devices do not exhibit polarisation switching (PS) with applied bias current.

3.2.3 Experimental methods

To experimentally operate VCSELs as high-speed, spiking photonic neurons, the principle of injection locking of these devices to an external coherent optical source is utilised. In the reported experiments, the external injection signal is provided by a tuneable laser source. This permits for precise tuning of the externally injected optical signal to match the VCSEL's resonance.

The information input to the VCSEL is realised via external modulation of the optically injected signal. Various degrees of freedom can be used to modulate the optical injection signal, including phase and intensity, with injection power, rate of frequency detuning [252] and polarisation [253] also playing a role. For the experiments covered in this thesis, only intensity modulation schemes have been utilised.

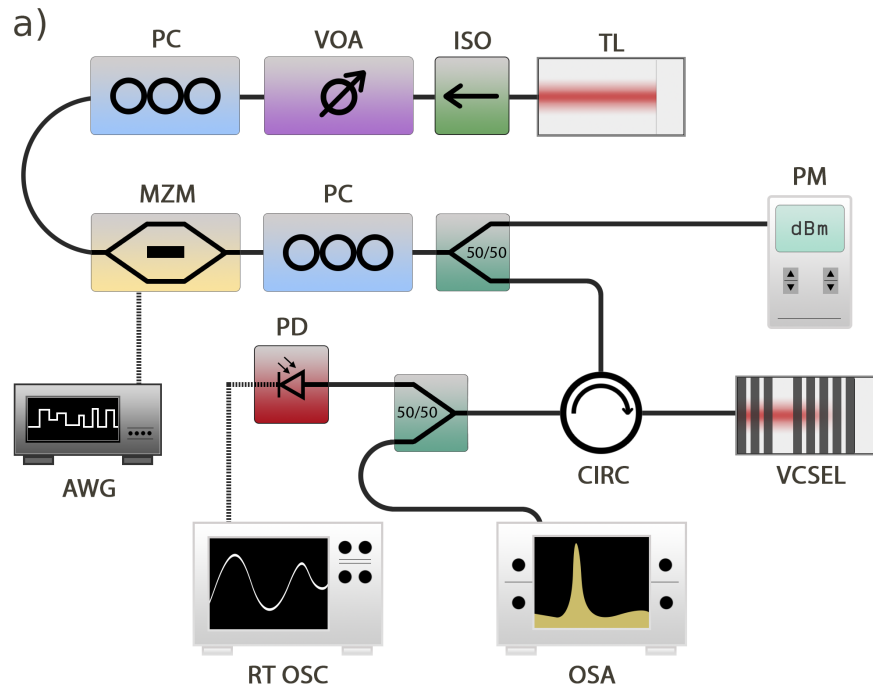


Figure 3.2: Diagram of the experimental setup used to investigate the VCSEL-neuron. TL – tunable laser; ISO – optical isolator; VOA – variable optical attenuator; PC – polarisation controller; 50/50 – optical splitter; PM – optical power meter; CIRC – optical circulator; OSA – optical spectrum analyzer; PD – photodetector; RT OSC – real-time oscilloscope; AWG – arbitrary waveform generator. Reproduced from [254].

The experimental setup used in this work is depicted in Fig. 3.2. This setup is realised using following functional blocks in respective order:

1. An **optical injection line branch** is required to produce the external optical injection signal. This is formed by a tuneable laser source, followed by an optical isolator (ISO) to ensure unidirectional coupling between the lasers. Following that, the signal is passed through a variable optical attenuator (VOA) for optical input power control. Then, the optical signal is polarisation matched to the modulator using a fibre-based polarisation controller (PC1). The amplitude modulation is performed using a MZM. The RF signals for the modulator are generated using an arbitrary waveform generator (AWG) operating at 12 GSa/s and incorporate an additional 10 dBm RF amplifier. After that, the modulated optical injection signal is polarisation matched to the desired mode of the VCSEL using a second

fibre-based polarisation controller (PC2). A portion of the optical input signal is acquired using a fibre-based power splitter (50:50 or 90:10) to monitor the optical injection power using a power meter (PM).

2. A **three-way signal pathway** is required, to feed the injection signal into the VCSEL, and to acquire the latter's response (output). This can be realised using an optical circulator (CIRC). In this component, signals are routed from ports 1 \rightarrow 2 (injection into the VCSEL), and from ports 2 \rightarrow 3 (collecting laser response).
3. An **output analysis** branch is required to analyse the VCSEL's optical output traces. Using a 50:50 optical fibre coupler, the signal is split for analysis with an optical spectrum analyser (OSA) and a fast real-time oscilloscope (RT OSC). In the oscilloscope line, the optical signal is O/E converted using a high-speed, fibre-coupled amplified photodetector (PD, Thorlabs PDA8GS).

The main focus of this thesis is on the analysis of excitable and spiking dynamics in photonic and optoelectronic systems, and these will constitute the main body of experiments described in the following sections. With respect to the frequency detuning between the injected optical signal and the VCSEL's resonance, the injection in the experiments shown here is typically operated with a weak negative detuning $\Delta\omega$ [220]. The operational point V_{MZM} (bias voltage) of the Mach-Zehnder modulator is set between the quadrature and the maximum of the output power transfer curve to achieve near-linear relation between input amplitudes and output light intensity modulations, and to preserve maximal achievable modulation amplitude. The modulations used to deterministically elicit optical spike events from the VCSEL-neurons typically take the form of square-shaped pulses of negative polarity (brief drops in optical power).

3.3 Characterisation of excitable spiking dynamics in the VCSEL-neuron

This section describes the experimental dynamical characterisation of the injection-locked VCSEL neuron. This includes experimental demonstration of refractory periods

in spike firing, spike onset latency characterisation as a function of input pulse amplitude, and sub-threshold pulse train integration. The measurements were performed using a 1310 nm commercially-sourced, fibre-pigtailed single transverse-mode VCSEL (RayCan). In these experiments, the VCSEL was biased with a current of $I_{\text{VCSEL}} = 5 \text{ mA}$, well above its lasing threshold current $I_{\text{thresh}} = 1.65 \text{ mA}$ (LI curve is shown in Fig. 3.3). Under such bias current and with a stabilised temperature of $T = 298 \text{ K}$, the VCSEL's optical spectrum (shown in Fig. 3.3) exhibited two distinct peaks that correspond to the two orthogonal polarisations of the fundamental transverse mode of the device. These are referred to as *parallel polarisation* (dominant peak, denoted λ_y) and *orthogonal polarisation* (secondary peak, denoted λ_x). The injection locking signal was provided by an O-band tuneable laser source (Santec TSL210V). This signal was injected into the orthogonally-polarised emission mode of the VCSEL.

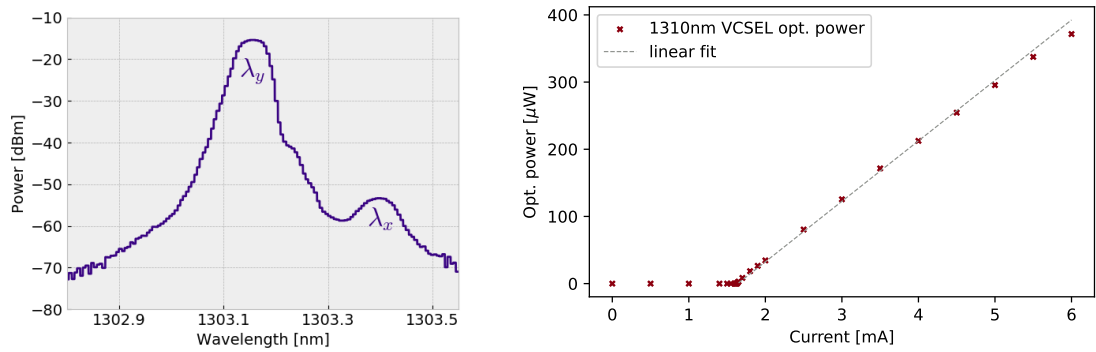


Figure 3.3: (Left) Optical spectrum of a free-running VCSEL biased with current of $I_{\text{VCSEL}} = 5 \text{ mA}$. The spectrum exhibits two polarisation modes referred to as λ_y (parallel) and λ_x (orthogonal). Reproduced from [255]. (Right) LI curve of the used 1310nm VCSEL.

3.3.1 Spike shape characterisation

By controllably perturbing the optical injection signal under a specific set of conditions, excitable spikes can be elicited in the VCSEL. To evaluate the character of these excitable responses, the injection into the VCSEL neuron was modulated (by the MZM) using an RF signal with rectangular-shaped power drops of $t_{\text{pulse}} = 0.5 \text{ ns}$. The injection power was $P_{\text{inj}} = 168 \mu\text{W}$ with detuning $\Delta\omega = -6.35 \text{ GHz}$. The modulated optical injection

signal is shown in Fig. 3.4(a). The responses from the VCSEL-neuron were collected through a polarising beam-splitter to separate contributions of both of the mutually orthogonal polarisations. These polarisation-filtered signals are shown respectively in Fig. 3.4(b,c), with the composite signal (sum) shown in Fig. 3.4(d). These time traces confirm how the excitable spikes VCSEL-neuron (here for orthogonal optical injection) emanate from interplay between the two polarised modes in the VCSEL. Further temporal analysis of the spike shape was performed using the composite output, and full-width at half-maximum (FWHM) spike width was calculated as $t_{\text{FWHM}} = 219 \pm 26$ ps (one standard deviation interval, calculated from $n = 35$ observed spikes).

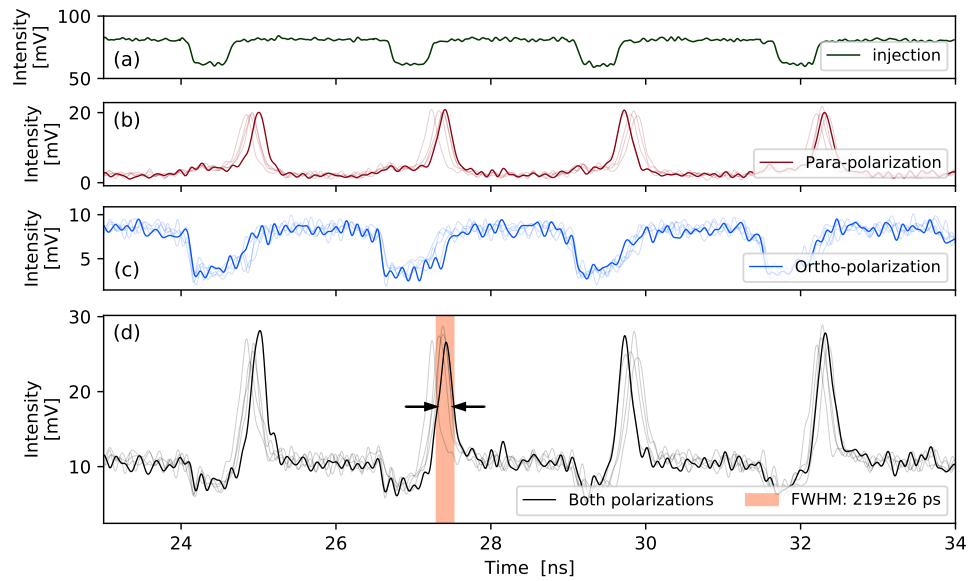


Figure 3.4: Excitable spiking in the VCSEL neuron under orthogonally-polarised optical injection. (a) Optical input signal with the encoded modulated perturbations to elicit spike events. (b) Parallel and (c) orthogonal polarisation-filtered signal collected from the output of the modulated VCSEL-neuron. (d) Composite signal of both polarisations, as typically recorded from the VCSEL-neuron output. The estimated FWHM of the output neuron-like excitable spikes is reported.

3.3.2 Refractory period characterisation

To establish the temporal timescales at which excitable dynamics can be elicited in the injection-locked VCSEL-neuron, a refractory period test was performed. This was done by utilising a series of triggering optical perturbation doublets with gradually

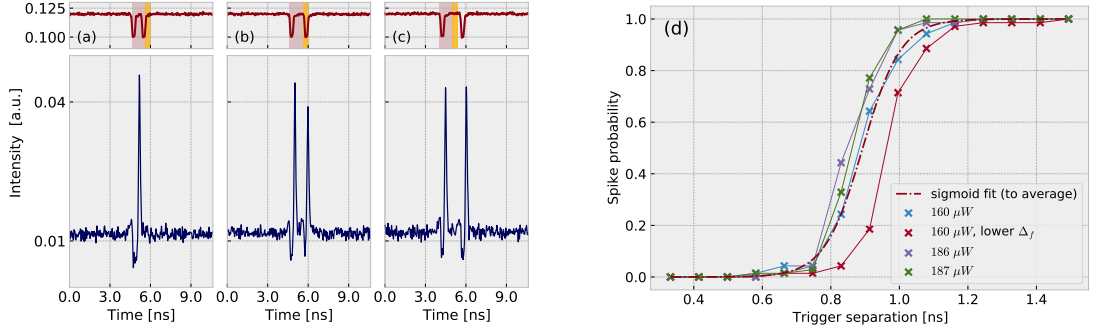


Figure 3.5: Refractory period measurement in the VCSEL-neuron. Upper (red) time traces show pairs of input perturbations with increasing temporal separation. The second perturbation arrives (a) within the absolute refractory period (red shading), (b) within the relative refractory period (orange shading), (c) after the refractory period, therefore eliciting a second spike. (d) Probability of firing a second optical spike versus temporal separation between input perturbations over $n = 70$ consecutive measurement cycles. Reproduced from [255].

increasing pulse separation. The perturbation waveform provided to the MZM consisted of rectangular-shaped power drops of $t_{\text{pulse}} = 249$ ps. The pulse separation t_{sep} was gradually increased from 332 ps to 1.494 ns with increments of 83 ps. Both the input modulation traces and the VCSEL-neuron's optical output time traces are shown in Fig. 3.5(a)-(c).

In each individual input perturbation doublet, the first perturbation always triggers an optical spiking response since it arrives when the VCSEL-neuron is in a locked (resting/quiescent) state. Firstly, in the case shown in Fig. 3.5(a), the second perturbation arrives $t_{\text{sep}} = 750$ ps after the first. In this case, the dynamical system is still within the so-called absolute refractory period ($t_{\text{sep}} < T_{\text{abs}}$, highlighted as a red shaded area in Fig. 3.5(a-c)) and cannot elicit a second optical spiking response. Once the temporal separation between the two perturbation reaches the edge of the absolute refractory period $T_{\text{abs}} < t_{\text{sep}} \approx 830$ ps, occurrences of second optical spikes can be observed. Fig. 3.5(b) shows the case for $t_{\text{sep}} = 1.08$ ns. For perturbations arriving within the relative refractory period window ($T_{\text{abs}} < t_{\text{sep}} < T_{\text{ref}}$), the probability of spiking is reduced and the optical spike amplitude is slightly decreased (as the system can be considered not fully recovered to elicit a full-amplitude spike). Finally, perturbation

doublets arriving with separations higher than the relative refractory period ($T_{\text{ref}} < t_{\text{sep}}$, such as $t_{\text{sep}} = 1.49$ ns, Fig. 3.5(c)) elicit the second optical spiking response both with high probability and full amplitude.

Fig. 3.5(d) plots the second optical spike firing probability in the VCSEL-neuron as a function of trigger doublet temporal separation. Results are plotted for various cases of optical injection powers with a mean sigmoid fit provided over all the recorded measurements. The optical injection power alters the locking strength between the injection signal and the VCSEL, which in turn influences the minimum required input perturbation strength required to elicit optical spikes. Fig. 3.5(d) shows a clear transition between absolute (no spiking observed) and relative refractory period (spiking occurs, but with reduced probability) followed by the fully recovered state (spiking probability very close to unity). Therefore, to ensure highly reliable activation of optical spikes with precise timing, the VCSEL-neuron should be operated with minimal inter-spike temporal separation equal or larger to the relative refractory period.

3.3.3 Optical spiking trigger latency evaluation

Another aspect that influences the character of the VCSEL-neuron spiking responses is the amplitude (strength) of the input perturbation. In excitable dynamical systems such as the photonic VCSEL-neuron, only input perturbations of sufficient strength that reach (or exceed) the spike activation threshold elicit a spike firing event. Furthermore, the shape of the spike is defined by the system, and is independent of the perturbation shape. Previous work has shown that optical spike onset latency in excitable micropillar lasers is influenced by the perturbation energy [256], representing a natural, nonlinear input encoding mechanism in the temporal domain.

To investigate the existence of this phenomena in the spiking VCSEL-neuron, a sequence of super-threshold optical pulses with gradually increasing strength is used as injection (input) modulation. The optical injection power was set at $P_{inj} = 162$ μ W and the injection frequency detuning at $\Delta\omega = -7.2$ GHz (with respect to the VCSEL's orthogonally polarised mode). By adjusting the injection strength and polarisation of the optical input signal, the VCSEL-neuron was brought close to the excitable threshold,

such that even small perturbations would be sufficient to elicit spikes. The input modulation consisted of $n_{pulses} = 9$ rectangular-shaped negative optical pulses (power drops), each with a temporal length of 500 ps and separated by 2 ns. The drop amplitude was gradually increased from 20% to 100% of the AWG digital-to-analog converter (DAC) range of 1 V peak-to-peak (p-p), using fixed increments to create equidistant levels while preserving the super-threshold character of the pulses. The last, strongest perturbation was used as a local timing reference, and all the follow-up spikes were timed against this last, guaranteed spike event in the system.

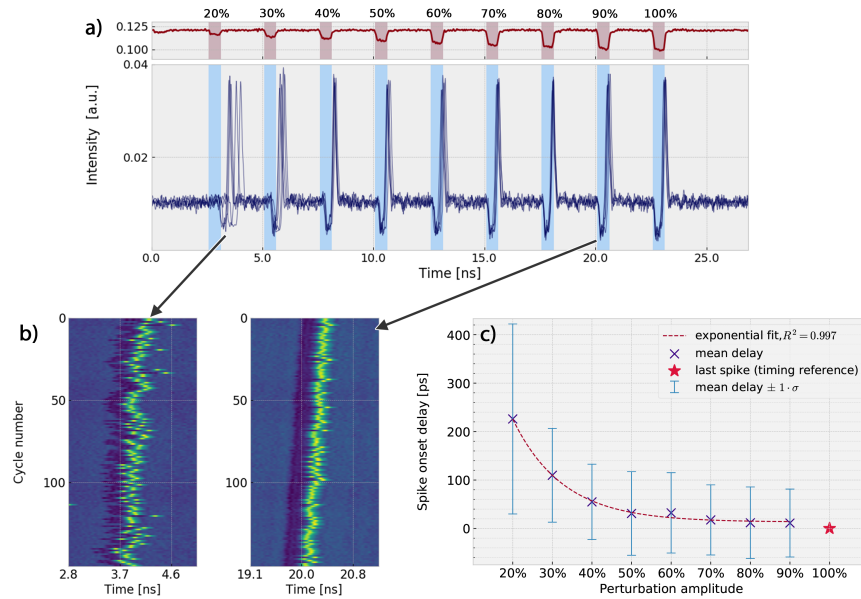


Figure 3.6: Demonstration of optical spiking latency in the VCSEL-neuron. (a) Variation in spike onset latency and trigger irregularity (jitter) as function of trigger amplitude. Overlay of $n = 5$ measurements. (b) Temporal maps of $n = 150$ measurement cycles, showing responses to 20% (left) and 90% (right) strength perturbations. (c) Mean spike onset delay time as a function of perturbation strength. Total of $n = 150$ repetitions, the last spike is used as timing reference. Reproduced from [255].

Fig. 3.6 demonstrates the effect of perturbation strength on the excitable spike activation latency in the VCSEL-neuron. Fig. 3.6(a) bottom (blue traces) shows an overlay of $n = 5$ superimposed output time traces recorded from the VCSEL-neuron when subject to the injection of the modulated optical signal with encoded perturbations depicted in the upper (red) trace. These five measurements were recorded consecutively.

All the perturbations are in a form of optical power drops, and all are of sufficient

strength to elicit a single spiking response from the VCSEL-neuron. For the weakest (left-most) super-threshold perturbations (denoted as having 20% amplitude), the produced spiking response exhibits both a pronounced onset delay (as depicted by the expected activation window shown in background blue shading) and a notable irregularity of activation times (jittering). It is worth noting that all the recorded spiking responses exhibit the same amplitude independent of trigger amplitude. As the perturbation strength increases, the jittering effect and activation irregularity settle down, and the spike triggering process becomes very well controllable, allowing for precise spike timing.

To provide further means for evaluation of triggering stability in both edge cases (weakest and near strongest perturbation), the spikes elicited in these two cases across all the recorded time traces ($n = 150$) were rendered in two focused temporal maps shown in Fig. 3.6(b). In these maps, each row represents a single time trace, with signal amplitude shown in colour, where green colour represents high output signal amplitude (spike crests) while blue represents low amplitude (background, no activity). Both of the shown maps demonstrate how the perturbation resulted in optical spike activation in 100% of the recorded cases, with the spikes triggered by the low-amplitude perturbation exhibiting pronounced temporal jittering (that vanishes for higher amplitude perturbations). The slight diagonal drift of the spiking events shown over subsequent measurements is not a feature of the measurements, but is introduced in the data processing (temporal map rendering) step. The presence of negative values within some of the error bars results from the timing reference being one of the spikes (that also experiences jittering).

Finally, Fig. 3.6(c) provides data on the average spike activation delay as a function of the perturbation strength. The spike elicited by the strongest (last) perturbation is used as a timing reference and its delay is therefore set to zero. Statistical evaluation of the onset delay was performed over the complete set of $n = 150$ recordings. For 20% amplitude optical perturbation, we observe mean activation delay of $\overline{t_\Delta} = 226$ ps, with jittering captured in the error bar showing single standard deviation interval $\sigma_t = 202$ ps. Notably, the activation timing irregularity rapidly settles down ($\sigma_t = 98$ ps for 30% amplitude perturbation, $\sigma_t = 72$ ps for 100% amplitude perturbation). The observed onset delay gradually decreases towards $\overline{t_\Delta} = 12$ ps for the strongest (100% amplitude)

optical perturbation, which is near the oscilloscope acquisition precision limit. This data confirm the previously observed behaviour, with both the mean spike onset delay and the delay standard deviation (due to the jittering) being increased for low amplitude trigger perturbations, and stabilisation observed for stronger perturbations.

In summary, these observations hint towards the capability of the VCSEL-neuron to encode input stimuli and information into fast optical spikes by means of temporal latency coding. A similar encoding principle was observed in, among others, spiking neurons in the retina, connecting it with the visual cortex in the brain [257].

3.3.4 Demonstration of leaky integrate-and-fire functionality

In the field of computational neuroscience, leaky integrate-and-fire (LIF) models represent one of the simplest and most widely used spiking neuronal models (see Sec. 1.2.4). LIF models exhibit a short term memory, represented by an exponentially decaying state variable. Since memory is one of the key functionalities for information processing, a LIF neuron can be used to build a stateful computing system.

Here, the leaky integrate-and-fire functionality was evaluated in the VCSEL-neuron for sequences of sub-threshold optical input perturbations. This was demonstrated using a long wavelength, 1310 nm VCSEL biased well above lasing threshold with current at $I_{\text{VCSEL}} = 5.0$ mA and constant temperature at $T = 298$ K. Injection locking was performed by subjecting the device to optical input signals with 183 μ W of optical power and injection frequency detuning $\Delta\omega = -6.0$ GHz (further from spike firing threshold) or $\Delta\omega = -6.35$ GHz (closer to spike firing threshold). The results from this experiment were published, among other dynamical characterisation procedures, in [236]. The input waveform includes four individual time-multiplexed optical perturbation patterns, including: first, a single super-threshold input pulse to showcase the excitability threshold in the system. Following this initial strong perturbation, there are three patterns of sub-threshold, return-to-zero optical input pulses ($t_{\text{pulse}} = 83$ ps, 50% duty cycle). These patterns include a single, a doublet and a triplet of sub-threshold pulses, each pattern separated by $t_{\text{sep}} = 2$ ns.

The results are shown in Fig. 3.7 for two different initial injection frequency detunings

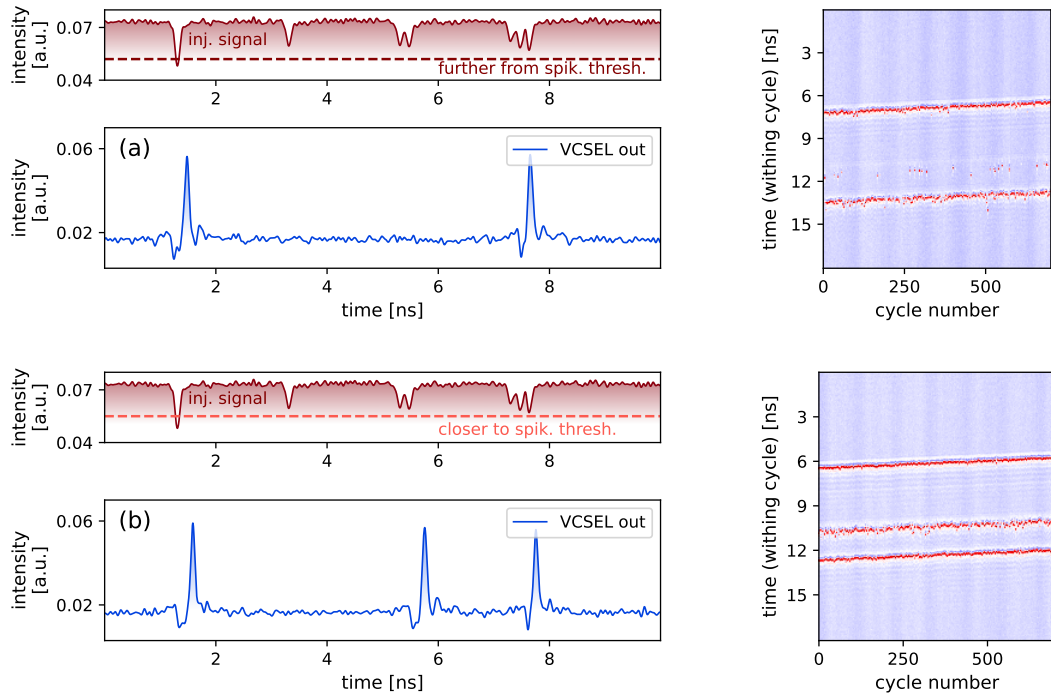


Figure 3.7: Demonstration of integrate-and-fire spiking responses in the VCSEL-neuron. The VCSEL-neuron is operated (a) further, (b) closer to the excitability threshold by adjusting injection frequency detuning. Red traces represent an example of the modulated injected optical signal (same demonstrative trace shown in both figures), while blue represent responses from the VCSEL-neuron.; Temporal maps constructed over $n = 750$ repetitions that plot the continuous time traces with the spiking and quiescent responses from the VCSEL-neuron represented in red and blue colours, respectively.

that correspond to the cases where the steady state is respectively further (Fig. 3.7(a)) and closer (Fig. 3.7(b)) to the device’s excitability threshold. In both cases, the red traces show the input optical modulation, with the approximate threshold level depicted by a dashed line. The output optical trace from the VCSEL-neuron is shown in blue. In either case, the first super-threshold input perturbation successfully elicits an optical spike. Without any short-term memory functionality, any of the sub-threshold perturbations would be insufficient to trigger any response. However, both the single traces and temporal maps (based on $n = 750$ repeated recordings) clearly highlight the presence of a spike firing event in the case of the triplet in Fig. 3.7(a) and for both triplet and doublet in Fig. 3.7(b). Such reliable triggering can only be achieved through a short-term memory effect like (leaky) integrate-and-fire. Therefore, the presence of

this mechanism is confirmed in the injection-locked photonic VCSEL-neuron. This LIF feature enabled the experimental demonstration of the high-speed pattern recognition task reported in [236].

3.4 Experimental information encoding

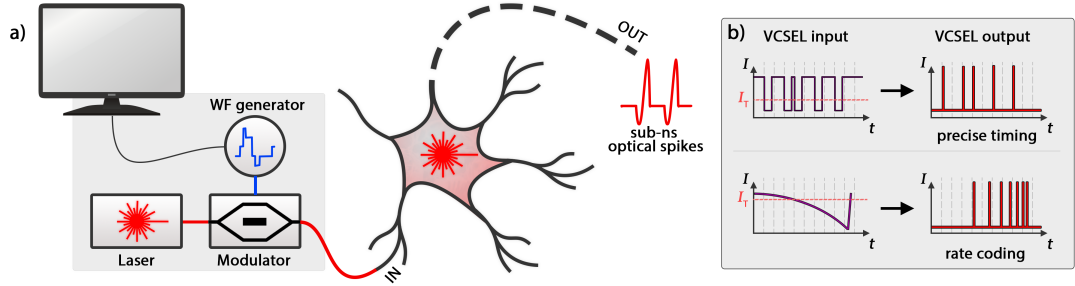


Figure 3.8: a) Simplified schematic demonstrating the operation of a photonic VCSEL-neuron. b) Working principle of the two distinct information encoding regimes in the VCSEL-neuron: precise temporal coding and rate coding.

In this section, multiple optical spike-based information encoding mechanisms are demonstrated in the photonic VCSEL-neuron. The results reported in this Section have been published in [255]. In general, we can differentiate between two main types of optical input signals to modulate the VCSEL-neuron: *digital* (binary) and *analog* (continuous) inputs, as depicted in Fig. 3.8(b). For each of these two general cases, the spiking responses from the device will exhibit a unique character, enabling the use of VCSEL-neurons for spike-encoding information schemes. As in the previous cases, the external injection signal was polarisation matched to that of the VCSEL's orthogonally polarised mode, and was configured with an optical input power ranging from $P_{inj} = 160 \mu\text{W}$ to $180 \mu\text{W}$ and set with a frequency detuning $\Delta\omega = -6.7 \text{ GHz}$.

3.4.1 Optical digital-to-spike (DtS) format conversion

First, a temporally precise optical digital-to-spike (DtS) information encoding mechanism is demonstrated in the VCSEL-neuron. Here, the input data is represented as a sequence of binary pulses. This data stream is directly encoded within the VCSEL-neuron

into a distinct spike train in the optical domain, utilising the all-or-nothing excitable thresholding functionality of the photonic VCSEL-neuron. To operate in this regime, the temporal spacing between individual trigger pulses needs to be above the system's absolute refractory period (Sec. 3.3.2), and the input binary pulses need to be of sufficient strength to exceed the spiking threshold and elicit spikes with high reliability and low temporal jittering.

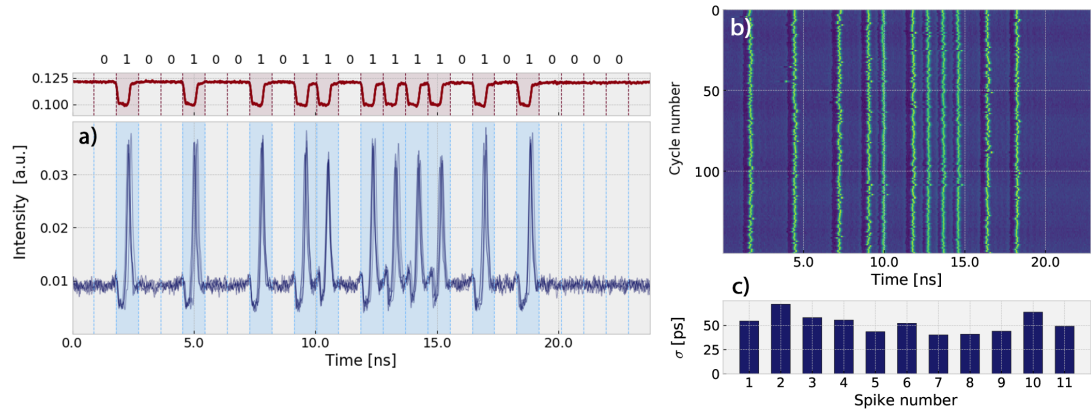


Figure 3.9: Digital-to-spike information encoding in the VCSEL-neuron for return-to-zero (RZ) input code. (a) (top, in red) Input sequence of 24 bits representing in binary the word "IoP", an acronym of the Institute of Photonics; (bottom, in blue) superposition of five time series recorded at the output of the VCSEL-neuron. (b) Temporal map of $n = 150$ consecutive repetitions of time traces recorded from VCSEL-neuron. (c) Standard deviations of timing for each spike across n repetitions. Reproduced from [255].

Fig. 3.9 shows the experimental results for the DtS encoding in the VCSEL-neuron. In DtS coding, digital signals that depend on clocking signals are converted to excitable spikes that are represented continuously in time. Fig. 3.9(a) shows both the input modulation waveform (in red, upper plot) and the response (in blue, bottom). The input data sequence consists of 24 individual binary values initially encoded as square-shaped optical power drops using an unipolar, return-to-zero (RZ) scheme with bit length of 916 ps ($t_{\text{ON}} = 498$ ps, $t_{\text{OFF}} = 418$ ps), corresponding to an input bit rate of 1.09 Gbps. The binary sequence used in this experiment represents the three-byte word "IoP" (acronym of the Institute of Photonics at the University of Strathclyde) represented in binary ASCII form of 24 bits. The binary representation is shown as a series of ones and zeros at the top of the figure. The optical injection power was set at $P_{inj} = 178$ μ W

with an initial frequency detuning of $\Delta\omega = -6.7$ GHz. The AWG output voltage, RF amplification and MZM operation point were set to achieve approx. 18% amplitude drops for each perturbation drop ($K_p = 0.18$ following the notation defined in [231]). Bottom (blue) plots in Fig. 3.9(a) show an overlay of five time series captured at the optical output of the VCSEL-neuron when subject to the RZ-coded optical injection modulation. Multiple readouts were acquired to confirm the reliability and stability of the spiking process. The temporal map shown in Fig. 3.9(b) depicts a total of $n = 150$ repeated measurements from the VCSEL-neuron for the same input. Both the individual traces and the temporal map of the experiments confirm that the encoded input bit sequence elicits a corresponding, temporally-encoded spike train in the photonic neuron with an excellent degree of reproducibility. For each binary '1' input, a single spike is reliably elicited, whilst for '0' input bits, the VCSEL-neuron remains quiescent. No outliers in spike timing are further confirmed by the single standard deviation values plotted in Fig. 3.9(c). The achieved firing rate of 1.09×10^9 spikes/s exceeds the maximum firing rate predicted on other photonic neuronal models [258]. Furthermore, the selected word length of 3 bytes (24 bits) is not imposed by limitations in the system, and was selected for practical purpose of demonstration of the DtS functionality. In practice, using the principle of precise spike time encoding, the system can process arbitrarily long bit sequences without a loss of generality. In summary, these results demonstrate how precisely-timed spike trains can be elicited in the VCSEL-neuron with repetition rates of > 1 GHz, with very high repeatability and consistent, precise timing.

3.4.2 Continuous signal encoding

Furthermore, the signalling functionality of the photonic VCSEL-neuron goes beyond elicitation of single excitable spikes via input thresholding. Certain types of dynamical models and dynamical neurons exhibit a wide, rich array of dynamical behaviours, including (among others) continuous spiking, bursting etc (see Sec. 1.2.3). The injection-locked VCSEL-neuron exhibits a SNIC bifurcation, and therefore displays a behaviour similar to Class 1 excitable neurons [42]. This allows for use of analog input signals to elicit spikes in a continuous fashion, where local firing rate within the elicited spike

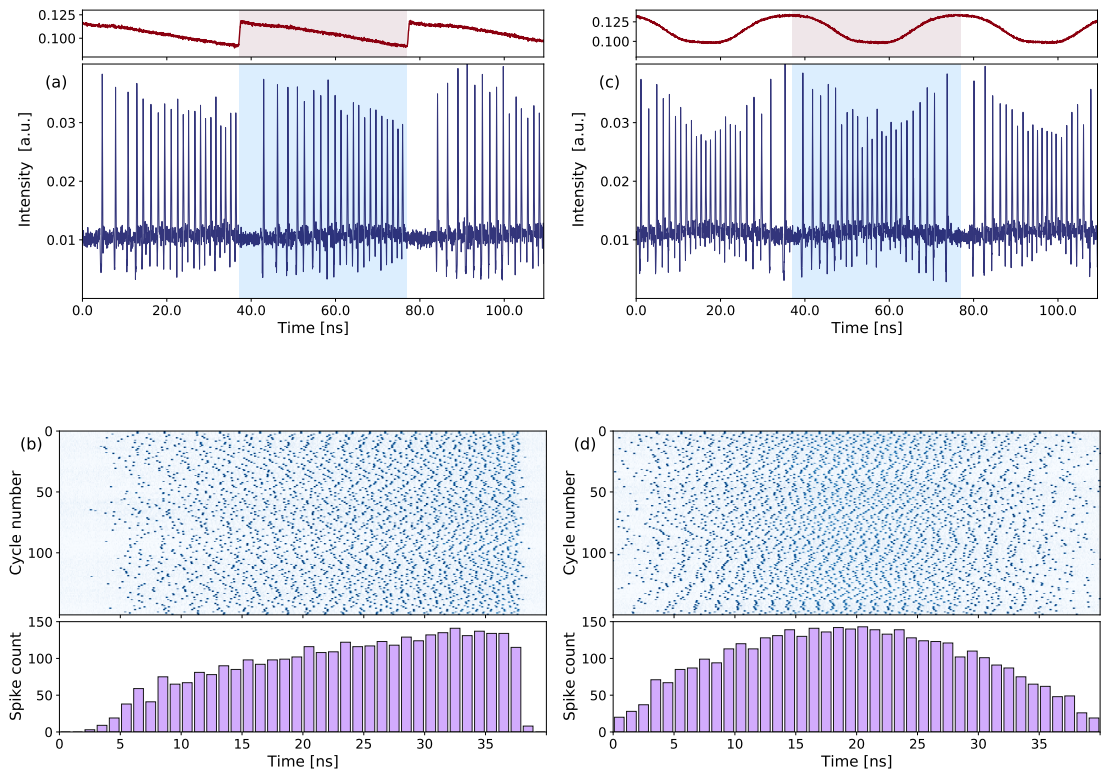


Figure 3.10: Continuous spiking rate-coding in a VCSEL-neuron, using (a,b) a sawtooth wave modulation, and (c,d) sinusoidal modulation. (a,c) Single traces for each modulation case, with modulation shown in red and VCSEL-neuron response in blue. (b,d) Temporal maps over $n = 150$ measurements, with histograms (1 ns bins).

trains directly corresponds to the strength of the input perturbations (i.e., the spiking period is directly proportional to the distance from the bifurcation point [259]). This permits us to effectively realise an all-optical, analog-to-firing rate information encoding scheme. Such an analog mode of operation is a fundamentally different behaviour than conventional analog-to-digital converters (ADCs) which convert analog signals into a digital representation. This behaviour will be investigated in the VCSEL-neurons using simple analog waveforms. Later, the spike firing rate encoding will be utilised to convert digital images into a purely temporal spiking representation.

Fig. 3.10 demonstrates the all-optical rate coding functionality of the VCSEL-neuron using two different analog modulation signals: a periodically monotonous sawtooth-shaped signal (left) and a non-monotonous sine-shaped signal (right). In both cases, the

signal frequency was set as $f_{\text{MOD}} = 25$ MHz. As the optical power of the injected signal decreases via modulating with the RF waveform, the system is pushed further from the stable (injection locked) state and the firing rate of the continuous spiking effect gradually increases. Continuous changes in injection power can therefore produce spike trains which directly encode the injection signal (stimuli) power in a fully temporal, neuromimetic fashion.

Fig. 3.10(a) demonstrates the encoding of a sawtooth-shaped modulation. As the injection power gradually linearly decreases from the initial fully injection locked steady state, a threshold is crossed and continuous optical spiking is activated at the VCSEL-neuron's output. These spikes are initially elicited in very low numbers, and their occurrence rate rapidly increases in as the power drops further. The full modulation range utilises an optical power drop of approx. 23% from the initial steady locking $P_{inj} = 143$ μ W to reduced optical injection at $P_{inj} = 110$ μ W. During this, the maximum achieved firing rate achieved is approx. 800 MHz. Fig. 3.10(b) captures the monotonic increase in spike firing rate both in the density increases within the temporal map and the increasing spike counts within 1 ns bins over a total of $n = 150$ repetitions.

This continuous, rate-coded spiking functionality is also demonstrated when a sine-wave modulation is used as the optical input signal in Fig. 3.10(c), utilising an optical power drop of approx. 25% from the steady locking ($P_{inj} = 143$ μ W) to $P_{inj} = 110$ μ W. The local spiking rate clearly periodically follows the locking signal modulation. Furthermore, the rate coding phenomena is visualised over $n = 150$ aligned repetitions in the temporal map shown in Fig. 3.10(d), together with a histogram of spike occurrences per each 1 ns long bin over all repetitions. Both the visual density in the temporal map and the histogram counts clearly follow the gradually increasing and decreasing character of the sine wave modulation.

3.4.3 Neuromimetic rate-coding of digital images

As previously discussed in Chapter 2, event-based vision systems provide a natural match to spike-based neuromorphic systems. Despite that, computer vision tasks and event-based processing of images remain a less explored territory within the domain of

light-powered spiking systems. In our recent manuscript [254], a single VCSEL-based spiking neuron has been proposed and demonstrated for operation mimicking that of a cone cell, a type of receptor in the eye that is sensitive to light intensity. The spiking VCSEL-neuron is an analog device, encoding data continuously in time using events (spikes) that are digital in amplitude.

In this experimental demonstration, a standard 8-bit value representing a spatial region (for example the R/G/B readout of a camera pixel or colour channel intensity of a digital image) is rate-encoded by the physical, spiking photonic VCSEL-neuron into a continuous train of optical spikes at GHz rates. This pixel readout (a real, bounded value) directly dictates the local spike firing rate of the continuous spike train. Furthermore, multiple pixels can be processed using a hardware friendly, single VCSEL-neuron arrangement by the means of time-division multiplexing (TDM), which significantly reduces the required experimental system complexity.

The experiment is realised following a procedure of multiple steps: 1) rearrangement of the digital image into a serialised vector in a computer, 2) creating a RF modulation waveform based on this vector using an arbitrary waveform generator (AWG), 3) modulating an optical signal with this waveform using a Mach-Zehnder modulator, 4) elicit spikes via injection locking of the VCSEL-neuron under the injection of the modulated optical signal.

Input data pre-processing for neuromimetic encoding

Initially, prior to feeding the input data into the VCSEL-neuron, the digital images (typically consisting of $3 \times N^2$ for RGB images without an alpha channel) need to be serialised and prepared for processing in the photonic system. The image is processed on a per-channel basis (R/G/B channels only), and the selected channel of size $N \times N$ is reshaped (flattened) into a vector of $n_{\text{total}} = N^2$ values. Arbitrary reshaping procedure can be utilised in this step, including row-by-row stitching or randomized serialisation.

Fig. 3.11 demonstrates the serialisation and pre-processing procedure for two different types of images, one containing a single channel (greyscale, Fig. 3.11(a)), and one containing two colour channels (red and blue, Fig. 3.11(d,g)). Each individual pixel

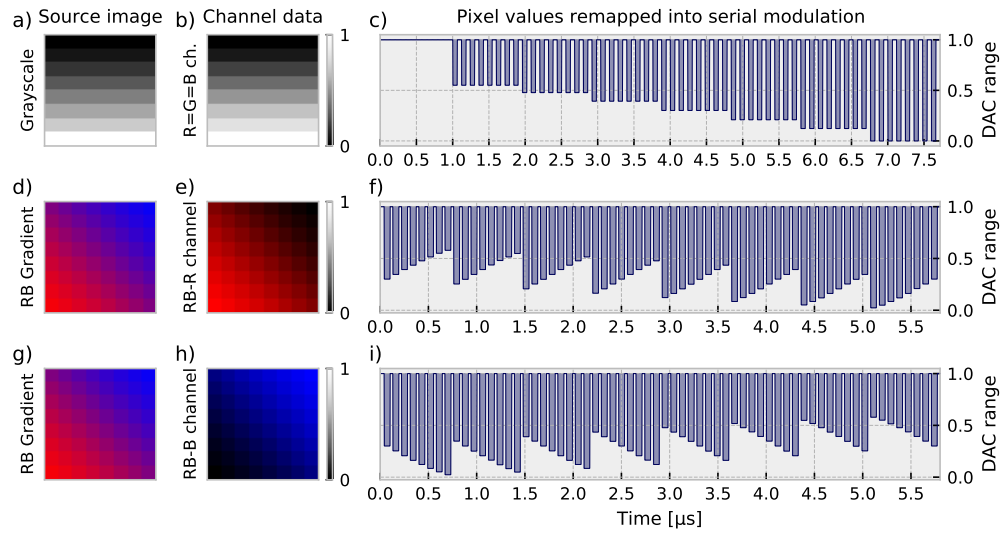


Figure 3.11: Digital image channel data pre-processing prior to rate-encoding in the VCSEL-neuron. Reproduced from [254].

value from the serialised vector is encoded (and time-domain multiplexed) into a square shaped, return-to-zero pulse of temporal length T_P , followed by zero period of length T_0 . A single return to zero pulse is referred to as a single *cycle* (total length $T = T_P + T_0$). In each square-shaped pulse, the corresponding pixel value P_n is encoded into the amplitude A_n using a linear relation with fixed $P_n = 0$ assigned for $A_n = 0$. For 8-bit values, the linear mapping relation can be described as:

$$A_n = \begin{cases} 0 & \text{if } P_n = 0 \\ \frac{1}{255} \cdot (1 - b) \cdot P_n + b & \text{if } P_n \neq 0 \end{cases} \quad (3.1)$$

where b is the baseline offset value of the linear mapping (for non-zero pixel values), here set as $b = 0.4$. Using this mapping, a fully injection locked state is obtained in the VCSEL-neuron for $P_n = 0$ (black) pixels, with spiking occurring for $P_n \neq 0$. Using this pre-processing, the single VCSEL can be considered as a single-pixel neuromorphic vision system studied over single cycle period T , exposed to constant light input of intensity proportional to P_n over interval $[0, T_P]$ and to darkness over $[T_P, T_P + T_0]$, with TDM utilised for processing of multiple pixels. The TDM step permits us to lower the amount of neurons required for processing, and could be in theory omitted if a sufficient

number of neurons can be operated in parallel (for example by the means of integrated neuromorphic VCSEL arrays [260] or even three-dimensional VCSEL neural network architectures [261]).

Operating the VCSEL-neuron for neuromimetic encoding

The operation of the VCSEL-neuron relies on injection locking of the VCSEL-neuron to an external (locking) optical signal. The RF modulation waveform created in the previous step is used as a modulation for this injection signal. The methods are described in Sec. 3.2.3, and the rate-coding functionality is realised using the same injection locking experimental setup as shown in Fig. 3.2. The experimental setup relies on fibre-optic components. Same as in previous cases, the VCSEL is a commercially available 1310 nm device (RayCan). The device was temperature stabilised at 298 K and operated at a bias current of $I = 5$ mA, well above the device's lasing threshold current ($I_{thr} = 2.9$ mA). Fig. 3.12 shows the lasing spectrum of this VCSEL, exhibiting a characteristic emission that corresponds to two orthogonal polarisations of the main transverse mode of the laser. The main lasing peak is referred to as *parallel* (λ_y) polarisation, with minor contribution from the *orthogonal* polarisation shown as a secondary peak at λ_x .

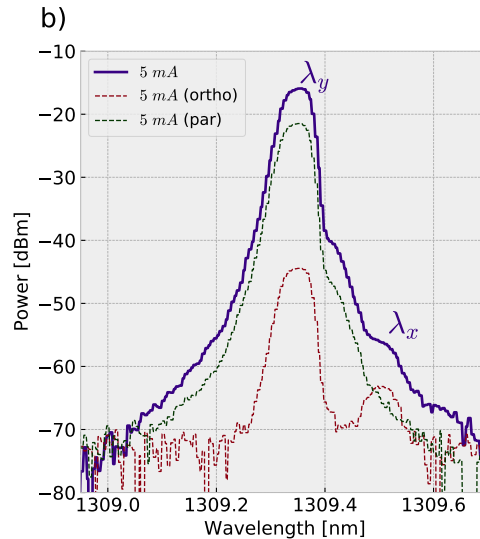


Figure 3.12: Emission spectrum of the VCSEL at $I = 5$ mA, with polarisation-selective filtering to demonstrate the two polarisations of the main transversal model. Reproduced from [254].

After eliciting the continuous spike trains via modulation of the injection into the VCSEL, the output optical time traces are collected on a real-time oscilloscope. All pixels in a single channel are acquired in a single-shot measurement on the oscilloscope. During each of the n_{total} recorded cycles, the spikes produced during the $[0, T_P]$ temporal window are counted, with spike occurrence being recorded based on crossing of a selected threshold value. This yields the spike counts c_n for the given cycle. The whole single-shot acquisition of pixels in one channel is repeated m times with the same set of conditions, yielding an average spike count value $\overline{c_n}$ for each pixel. Finally, the recorded spike count is converted into a mean, average spiking rate for each pixel as $r_n = \frac{\overline{c_n}}{T_P}$ and normalised using the maximum recorded spiking rate in the image $\max(r_n)$ as $\mathcal{P}_n = \frac{r_n}{\max(r_n)}$. In the final reconstructed image, the normalised rate values \mathcal{P}_n are assigned to their original positions in the image pixel array, creating a human-perceivable representation of the image based on the rate coded spike trains.

Rate-coding of gradient images

To demonstrate the rate-coding functionality offered by the VCSEL-neuron, two simple 8×8 pixel gradient images are first processed using the VCSEL system into TDM sequences of optical spike trains. These include a single-channel, grayscale (GS) image containing a simple vertical gradient linearly changing between white and black, and two-colour (two-channel) red-blue gradient, linearly changing between full blue and full red diagonally across the image. The images, their individual colour channel information and the respective RF modulation waveforms are shown in Fig. 3.11.

In the simplest case of the single-channel GS image (Fig. 3.11(a)), the modulation waveform is shown in Fig. 3.11(c). The waveform consists of negative-polarity pulses of temporal length $T_p = T_0 = 60$ ns with 50 : 50 return-to-zero duty cycle (total cycle length 120 ns), resulting in total encoding time 7.68 μ s. Following the formula shown in Eq. 3.1 and processing the primarily image left-to-right and then top-to-bottom, the first row of black pixels encodes into pulses with zero amplitude (no pulses). As the intensity increases in each row, the amplitude of the corresponding power drops follows. For this experiment, the injection power (with active RF modulation using a MZM) was

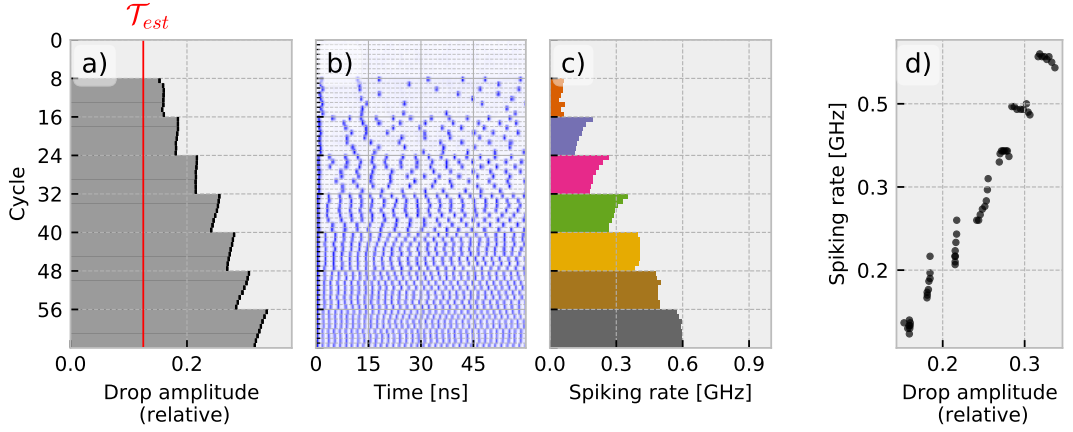


Figure 3.13: Spike-rate encoding of an 8×8 GS image with a VCSEL-neuron. (a) Output power levels from the MZM for each of the 64 power drops when modulating with waveform from Fig. 3.11(c); (b) Temporal maps created from the experimental data, showing the response of VCSEL-neuron during each modulation power drop of $T_p = 60$ ns (T_0 interleave period is omitted). (c) Mean spiking rates calculated for each cycle. (d) Spiking rate produced by the VCSEL-neuron as a function of input perturbation strength (drop amplitude). Reproduced from [254].

$P_{inj} = 182.3 \mu\text{W}$ and the initial frequency detuning of the external signal (with respect to the orthogonally-polarised mode of the VCSEL) was equal to $\Delta\omega = -6.29$ GHz.

The processed experimental output traces obtained for the GS image are shown in Fig. 3.13. Specifically, Fig. 3.13(a) shows the optical injection signal power levels (Mach-Zehnder output) for each of the 64 power drops when modulating with the waveform in Fig. 3.11(c). The slight skew of the produced power drops comes from the non-perfectly linear transfer function of the MZM and the limited bandwidth of the RF components used. Each of the power drops pushes the VCSEL-neuron beyond the excitability threshold (approximately highlighted as a solid red line in Fig. 3.11(a)) and into the continuous firing regime. The responses through all of the individual pixels are shown in the temporal map in Fig. 3.13(b), with major grid lines in the map separating individual rows (groups of eight pixels within the 8×8 pixel image). The GS image contains eight rows, and all pixels within a same row have an identical value - therefore, they should produce spike trains with equivalent firing rates at the output of the VCSEL-neuron. The rate is more significant than the local spike timing, since the spike trains exhibit some degree of local randomness (behaving as Poisson processes in

certain biological neurons [262]) and individual spikes within the ensemble may appear in different locations while preserving the repetition (rate) information. The gradually increasing density of observed spikes in the temporal map (Fig. 3.13(b)) and increase of the measured spiking rate (in Fig. 3.13(c)) confirm that the spike firing rate increases monotonically with the power drop amplitude (and therefore the input pixel intensity), in agreement with previous experiments [263].

A total of $n_{meas} = 3$ independent oscilloscope readouts were taken for the same GS image input under the same operating conditions to obtain the mean spiking rates shown in Fig. 3.13(c). These are highlighted with individual colours per each individual row of pixels in the image. Finally, Fig. 3.13(d) plots the direct relation between the amplitude of injected optical power drops for each pixel (the input image data), and the corresponding mean optical spiking rate from the VCSEL-neuron (the output). In certain pixels at row starts, a transient effect of locally higher-than-expected spiking rate was observed, with the effect gradually decreasing towards a stable, representative value. The transient occurred for pixels at (and near) row ends, suggesting that the stateful nature of the VCSEL-neuron coming from its LIF-character introduces more complex dynamics at transitions between pixel values. Overall, a good degree of consistency was observed among the independent readouts, confirming that the VCSEL-neuron can process input images reliably in a single shot fashion.

Furthermore, processing of multi-channel images is demonstrated on a diagonal red-blue (RB) gradient image shown in Fig. 3.11(d,g). The individual *RB-R* and *RB-B* channels are shown in Fig. 3.11(e) and Fig. 3.11(h) respectively, and the corresponding modulation waveforms of the RZ amplitude-coded square power drops are shown in Fig. 3.11(f) and Fig. 3.11(i) respectively. Each pixel intensity value was encoded into the power drop optical input amplitude using Eq. 3.1, temporal pulse length $T_p = 60$ ns with RZ period of $T_0 = 30$ ns, resulting in a total per-pixel processing time of $T_p + T_0 = 90$ ns and a total waveform length of 5.76 μ s per single channel. The frequency detuning $\Delta\omega$ was set as -5.59 GHz for both channels, and the injection power was set equal to $P_{injR} = 115.6$ μ W and $P_{injB} = 128.4$ μ W respectively for the RB-R and RB-B channels.

The processed experimental data of both RB-R and RB-B channels is shown in

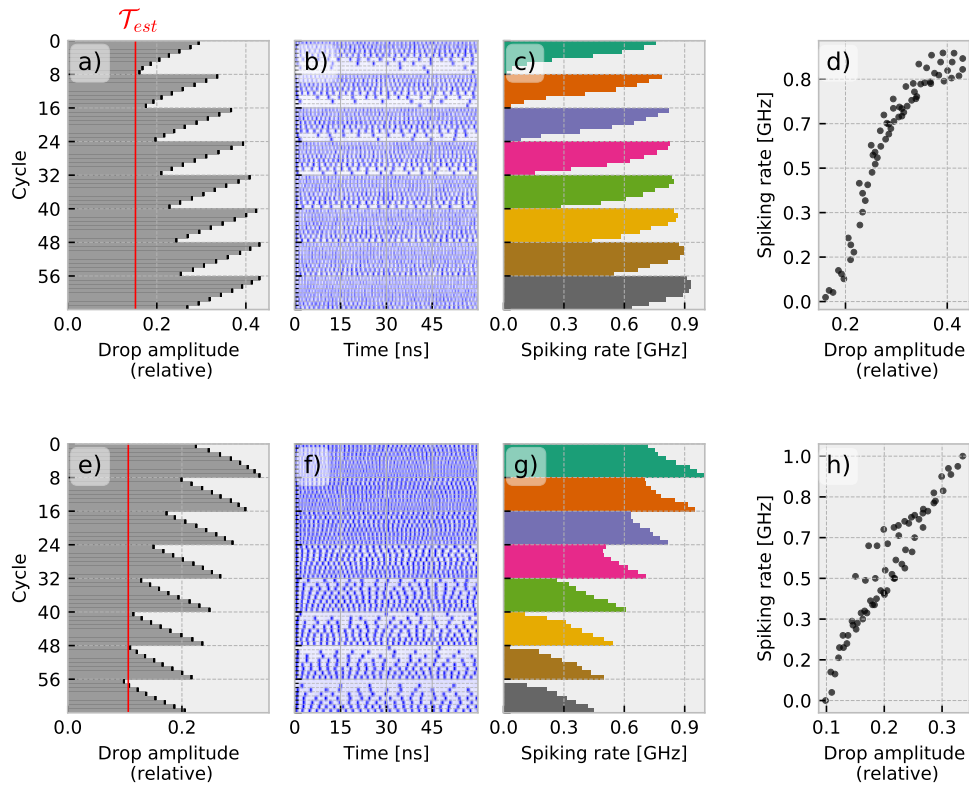


Figure 3.14: Spike-rate encoding of an 8×8 , two-channel RB image with a VCSEL-neuron. Results are presented for the individual RB-R (a-d) and RB-B (e-h) colour channels: (a, e) For each of the 64 pixels (cycle) in the RB-R (a) and RB-B channels (e) in the RB diagonal source image, the pixel intensity is encoded in the magnitude of the input perturbation (power drop) added to the optical signal injected into the VCSEL-neuron, with the estimated spiking threshold (\mathcal{T}_{est}) highlighted as a red line; (b, f) Temporal maps for the responses of the RB-R (b) and the RB-B channels (f); (c, g) Calculated average spiking rates for each cycle (pixel) in the RB-R (c) and RB-B channels (g); (d, h) Spiking rate produced by the VCSEL-neuron as a function of input perturbation strength (drop amplitude) in the RB-R (d) and RB-B channels (h). Reproduced from [254].

Fig. 3.14. Thanks to the gradually changing pixel values of the diagonal gradient, the waveform amplitudes created by appending values within a row (and then progressing on a row-by-row basis) resemble a saw-like shape. In either case, the source modulation Fig. 3.14(a,e) resulted in encoding of continuous spike trains at the output of the photonic VCSEL-neuron for RB-R and RB-B respectively. The time traces of the produced optical spike trains are depicted in the form of a temporal map in Fig. 3.14(b,f).

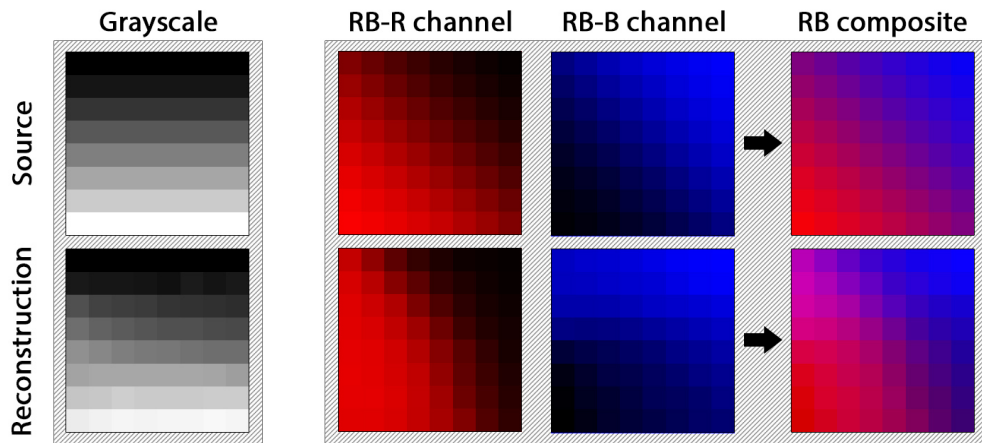


Figure 3.15: Source images and reconstructions based on the spike-encoded representation for GS and RB-R, RB-B channels. Source colour channels are shown in upper row, and channels directly reconstructed from the spike rates are shown in bottom row. Reproduced from [254]

Interestingly, recorded spike counts in Fig. 3.14(c,g) show again the transient effect of the short-term memory within the VCSEL-neuron, and how the character of the responses during sequential processing influences the result. In particular, in the RB-R channel Fig. 3.14(c), where the sequential processing produced much more pronounced transient differences between rows, there is an apparent over-firing effect with higher than expected spiking frequency. Meanwhile, the gentler high-to-low transitions in the RB-R channel Fig. 3.14(g) do not exhibit such pronounced effect. While this effect is not desirable in this particular encoding scheme, it might have interesting potential for advanced signalling that could accentuate and highlight rapidly changing events. Finally, the amplitude-to-rate function is depicted in Fig. 3.14(d,h) with saturation and higher degree of randomness observed in the transient effect influenced RB-R channel, and approximately linear responses observed in the RB-B channel.

Finally, a comparison is provided between the two source images (a total of three processed channels), and the direct reconstruction from their rate-coded spiking temporal representations produced by the VCSEL-neuron. This is shown in Fig. 3.15. The reconstruction of the original images was performed directly from the normalized spike firing rates (events/time) from the VCSEL-neuron, which are shown in Figs. Fig. 3.13(c) and Fig. 3.14(c,g) respectively. Source composite images and individual colour channels

are shown in the top row, while reconstructed composite images and channels are shown in the bottom row in Fig. 3.15. In all the image encoding procedures, both the reconstructed individual channels and composite images preserved the colour gradients well and exhibited good agreement between the source images and the reconstruction generated from the measured spike firing rates.

Rate-coding of a three-channel RGB image

Finally, to further demonstrate the practical performance of the rate-coding approach using the photonic spiking VCSEL-neuron, a larger-scale, 32×32 pixel, complex RGB image was used as a source for the system. To mitigate the effect of the transient overfiring observed in the previous RB-R image (Fig. 3.14(b)) that lead to locally increased optical spike firing rates, a randomised pixel processing order was generated and used for all the three individual colour channels of the full RGB image.

The optical input modulation waveform pulse sequence consisted of return-to-zero, $T_p = 53.3$ ns power drops whose amplitude encoded the pixel intensity value, and a $T_0 = 26.6$ ns zero interleave period. In total, each channel consists of 1024 pixels in a set randomised order, resulting in a total single channel processing time of $81.9 \mu\text{s}$. Due to the total waveform length and the greater number of processed pixels, temporal maps and bar plots of the obtained spike firing rates are not presented in this case.

The experimental parameters are $\Delta\omega = -4.9$ GHz for RGB-R; $\Delta\omega = -4.9$ GHz for RGB-G; $\Delta\omega = -5.25$ GHz for RGB-B, while the injected power was set to $P_{inj} = 120.9 \mu\text{W}$ for RGB-R; $P_{inj} = 119.6 \mu\text{W}$ for RGB-G; $P_{inj} = 121.5 \mu\text{W}$ for RGB-B.

Similarly to the previous cases of the smaller GS and colour gradient images, the pixels in the more complex RGB image are processed using a TDM scheme and injected serially into the spiking photonic VCSEL-neuron. In response, the laser produced rate-coded spike trains, and corresponding firing rates were calculated for each pixel. These are shown as a function of the corresponding pixel intensity values in each individual channel in Fig. 3.16 for red (R), green (G) and blue (B) channels respectively. These curves demonstrate the linear relation between the input (pixel intensity) and the output of the system (spike firing rate), with varying onset defined by the set of operational

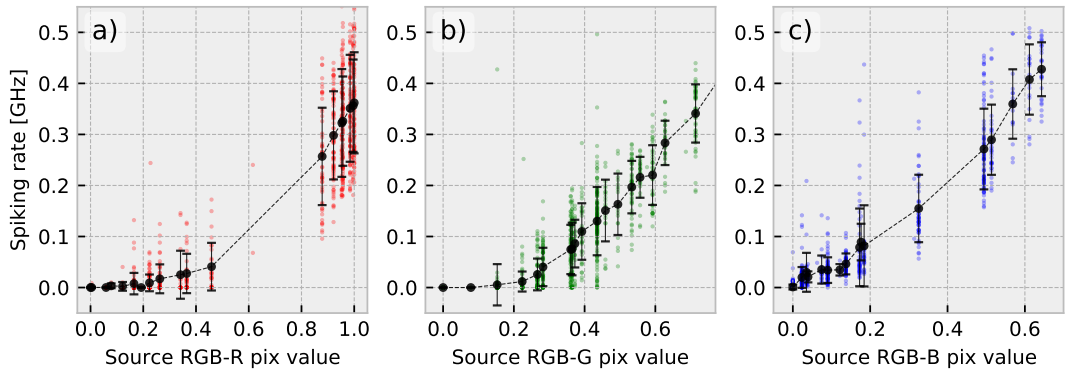


Figure 3.16: Mean obtained optical spike firing rates from the VCSEL-neuron as a function of the input pixel intensity values for each colour channel in the RGB image: a) red (R); b) green (G); c) blue (B). Each colour dot in the plots corresponds to an individually processed pixel. Mean values per each individual pixel value are shown as black dots, with error bars corresponding to single standard deviation. The sparsity observed in the red channel comes from the red colour distribution in the source image. Reproduced from [254].

conditions of the experiment. Due to the limited bit depth of the source pixel art image, there is a limited set of discrete, input pixel values in each channel, which allows for additional statistical processing of the firing rates using averaging (black dot) and standard deviations across each single-shot channel readout (black error bars). Error bars are provided only for pixels with spike occurrence counts of $n_{occ} > 3$. Finally, the dashed black curve connects the mean values, demonstrating a representative spike firing rate tone curve of the system.

The reconstruction of this larger, pixel-art style RGB image based on the rate-coded experimental optical spiking time traces is provided in Fig. 3.17 both for the individual colour channels and for the composite RGB image. The colour noise (spike firing rate variance) can be attributed to the transient overfiring effect occurring at certain pixels as they are randomly processed in the VCSEL-neuron. Overall, a very good degree of agreement is visible between the source and the reconstructed images, with the black point of some channels shifted in agreement with the tone curves shown in Fig. 3.16. It is important to highlight that the short-term memory, transient firing effect can be mainly attributed to the fact that a single VCSEL-neuron is operated in a TDM-based fashion, and a parallel processing via multiple VCSEL-neurons would very likely mitigate this

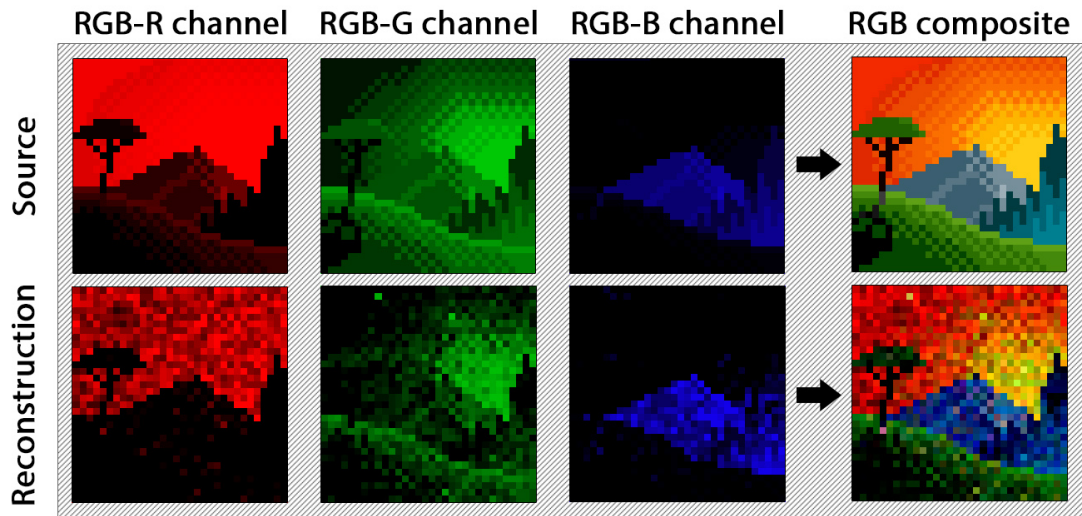


Figure 3.17: Processing of a complex, pixel-art 32×32 RGB image. (*upper row*) Individual R/G/B channels of the image, with source image shown at left. (*lower row*) Rate-encoded and reconstructed R/G/B channels produced by the VCSEL-neuron, with composite reconstruction image shown at left. The contrast of the individual reconstructed channels and directly reconstructed composite image was adjusted to better match to the source images and allow for clearer comparison. Reproduced from [254].

issue, especially for smoothly varying scenes (with respect to the GHz rates of processing in the system). Overall, the reconstructed RGB composite image demonstrates a good agreement with the source image, preserving the images key features.

3.5 Interfacing VCSEL-neurons with SiPh weight bank chips

The final section of this Chapter discusses and investigates experimentally the prospects of interfacing VCSEL-based photonic neurons with integrated photonic systems. Integrated photonic approaches allow for reduced circuit footprints and higher circuit density, both representing key metrics for practical computing platforms. Therefore, integrated neuromorphic photonic realisations are among the most promising solutions for practical applications.

A broad range of all-optical, neuron-like functionalities in VCSELs has been demon-

strated in VCSELs [242], proving them as viable candidates for high-speed optical spiking artificial neurons and spike-processing systems. However, a functional neuromorphic platform would require additional components beyond just the neurons themselves. At minimum, the required solutions may include i) interlinking and ii) signal weighting solutions. For these reasons, demonstrating the operation of VCSELs in conjunction with integrated photonic elements (able to offer both of these functionalities) provides an interesting and important research direction.

In this section, an experimental demonstration is provided for two different functional layouts that jointly utilise a single VCSEL-neuron in combination with a single silicon photonics (SiPh) microring resonator (MRR, representing part of an integrated photonic weight bank). The weight bank has been designed for operation primarily with WDM-based, continuous wave optical neural networking approaches. Therefore its prospects for processing high-speed (sub-nanosecond long) optical spiking signals remain unexplored to date. The results presented in this section are currently available as a pre-print [264], and are currently in the process of a peer-reviewed journal submission.

In the first functional layout, optical synaptic spike amplitude weighting is demonstrated in Sec. 3.5.1 for a VCSEL-neuron \rightarrow SiPh MRR arrangement (demonstrating a functionality that was previously theoretically proposed in [173]). In the second layout (Sec. 3.5.2), the optical spike rate coding mechanism (see description in Sec. 3.4.2) is demonstrated with the integrated platform. In this demonstration, a continuously firing VCSEL-neuron is directly modulated using an optical signal fed through the SiPh MRR (therefore utilising a SiPh MRR \rightarrow VCSEL-neuron arrangement). Using both of these uni-directional joint layouts, various functionalities are demonstrated that are of practical interest for realising functional, neuromorphic (spiking) photonic systems.

The SiPh chip was devised and fabricated by our collaborators, the team of Prof Prucnal at Princeton University (New Jersey, USA). The MRR forms a part of an integrated SiPh weight bank consisting of 4 MRRs [265] that allows for wavelength-division multiplexed broadcast and weight operation [173]. Here, all the MRRs are in the "add-drop" configuration with adjacent THRU/IN ports connected and adjacent ADD/DROP ports connected. The core layer consists of 220 nm thick Si, with 500 nm

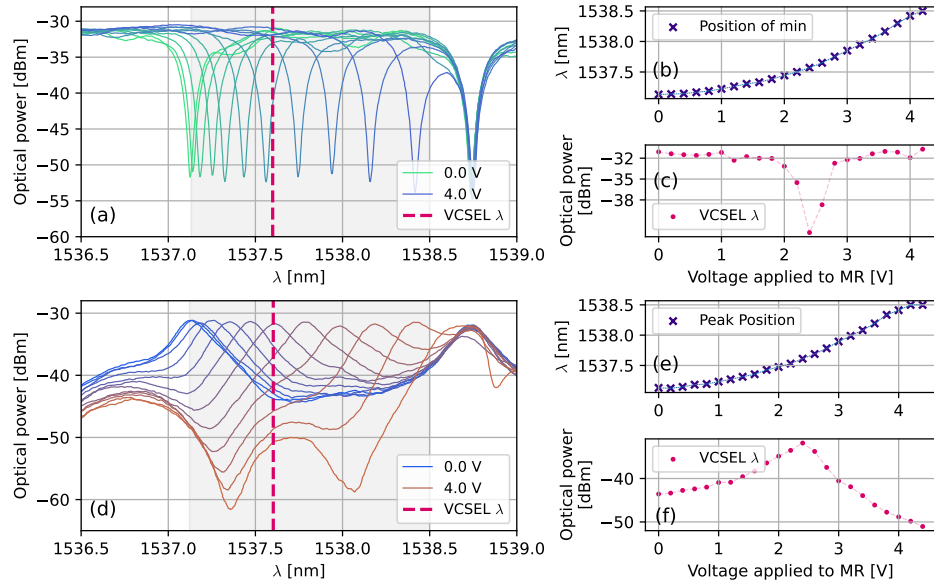


Figure 3.18: Characterisation of the MRR within the integrated SiPh weight bank. Upper three plots (a,b,c) correspond to characteristics recorded on the THRU port of the MRR, while (d,e,f) correspond to readout from the DROP port readouts of the MRR. (b,e) Wavelength of MRR resonance as a function of V_{MR} . (c,f) Recorded optical power at a fixed wavelength (dashed line) and maxima (minima) position as function of applied MRR bias V_{MR} .

wide bus, 200 nm bus-coupler gaps and ring waveguides with approx. 8 μm radius. Fig. 3.18 shows the spectra measured through the DROP and THRU ports spectra of a single selected MRR alongside the measured optical power at a given selected, fixed wavelength (which is highlighted by a dashed red line). Furthermore, the positions of a relevant extrema (peak or minimum) are given as a function of the bias voltage applied to the tuning element (heater) of the MRR for resonance tuning.

3.5.1 VCSEL-neuron spike weighting in SiPh MRR

The experimental setup for this demonstration is shown in Fig. 3.19. A commercially-sourced, packaged, fibre-pigtailed 1550 nm VCSEL was used. This VCSEL is operated at $I = 5 \text{ mA}$, well above the lasing threshold current of $I_T \approx 1.5 \text{ mA}$ (measured at 298 K). This VCSEL exhibited two polarised modes, with the main lasing mode being the one located at a longer wavelength (referred to here as a dominant lasing mode). The second

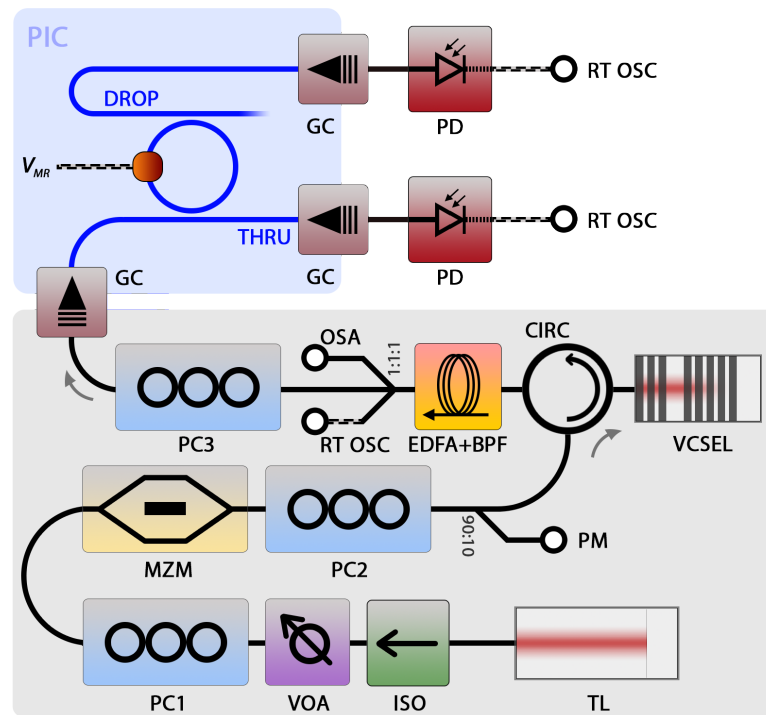


Figure 3.19: VCSEL-to-MRR experimental setup. TL - tuneable laser; ISO - isolator; VOA - variable optical attenuator; PC(1,2,3) - polarisation controllers; MZM - Mach-Zehnder modulators; PM - power meter; CIRC - circulator; VCSEL - vertical cavity surface emitting laser; EDFA - erbium-doped fibre amplifier; BPF - bandpass filter; RT OSC - real-time oscilloscope; OSA - optical spectrum analyzer; GC - grating coupler; PD - photodetector. Ratios for the optical couplers/splitters are indicated in the image.

mode, orthogonally polarised to the dominant lasing mode is referred to as a subsidiary mode. The wavelength separation between the two polarisation modes of the VCSEL was approximately 250 pm. The VCSEL emission spectra is shown in Fig. 3.20 both for a free-running operation (CW lasing) and during external injection-locking into the subsidiary (shorter wavelength) mode. Fig. 3.20 also shows that under external optical injection, switching between the two polarised modes of the VCSEL is obtained. The snapshot of the injection-locked spectrum was taken when operating the VCSEL-neuron in a spike firing regime; hence the widening shape observed in the two resonance modes. The laser was locked to the external optical injection from a remotely controlled tuneable laser source (EMCORE micro-ITLA). The RF modulation signal was generated by an AWG (Tektronix AWG7122B, 12 GSa s^{-1}) in the form of a train of super-threshold

activation perturbations (power drops). The RF modulation pulses were 750 ps (8 Sa) long, with 4.66 ns (8×7 Sa) spacing and 600 mV amplitude. This modulation ensured that the VCSEL-neuron operates as a source of deterministic, temporally equidistant optical sub-ns spikes. Prior to injection of the optical signal into the VCSEL-neuron, the

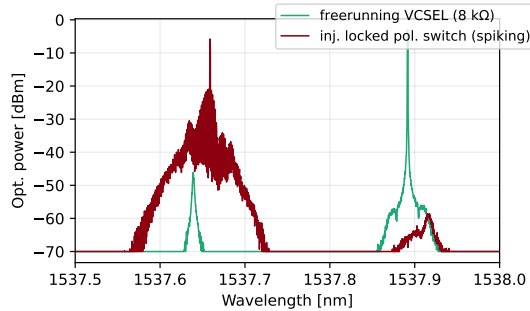


Figure 3.20: Optical spectra of the VCSEL when operated in (green) a free-running, CW lasing regime; (red) under the external injection of an optical signal into the device’s subsidiary mode, resulting in injection locking with associated polarisation switching. This injection locked trace was acquired while operating the device in a constantly spiking regime, hence widening the shape of the spectrum.

optical input path was split into two branches with a 90:10 fibre optic splitter. The 10% branch was used for external injection optical power monitoring, and the 90% power splitter branch was injected into the VCSEL through an optical isolator, with $\Delta\omega = -3$ GHz and $P_{inj} = 170$ μ W. The VCSEL-neuron output signal was collected through an optical circulator in the setup and amplified using an erbium-doped fibre amplifier (EDFA, PriTel) and equally split using a custom, 1-to-4 optical splitter (ThorLabs). One of the 1-to-4 splitter branches was used to provide signal into the SiPh chip, while another was used for optical spectrum monitoring using a high precision spectrum analyser (Apex Technologies). The third branch was used to monitor the time-dynamic responses using a real-time oscilloscope (with an additional photodetector to convert the optical signals into the electrical domain for analysis). The optical input and light collection into and out of the chip (THRU port, DROP port) were realised using vertical grating couplers. The optical weighting function of the MRR was performed by applying (sweeping) a DC signal V_{MR} and collecting the resulting output traces from the MRR for each applied DC bias value. Changing the DC voltage V_{MR} applied to the MRR heater

element permits us to tune its resonance, therefore allowing us to control precisely the optical output intensity exiting the device's THRU/DROP ports. This functionality permits us to control the attenuation of signals from these ports, therefore allowing us to weight the optical spiking signals generated by the VCSEL-neuron. Examples of measured time traces from the MRR's THRU and DROP ports for three different DC bias values are shown in Fig. 3.21. These demonstrate how the amplitude of a spiking pattern from the VCSEL-neuron can be effectively weighted by applying the bias voltage V_{MR} to the MRR heater element.

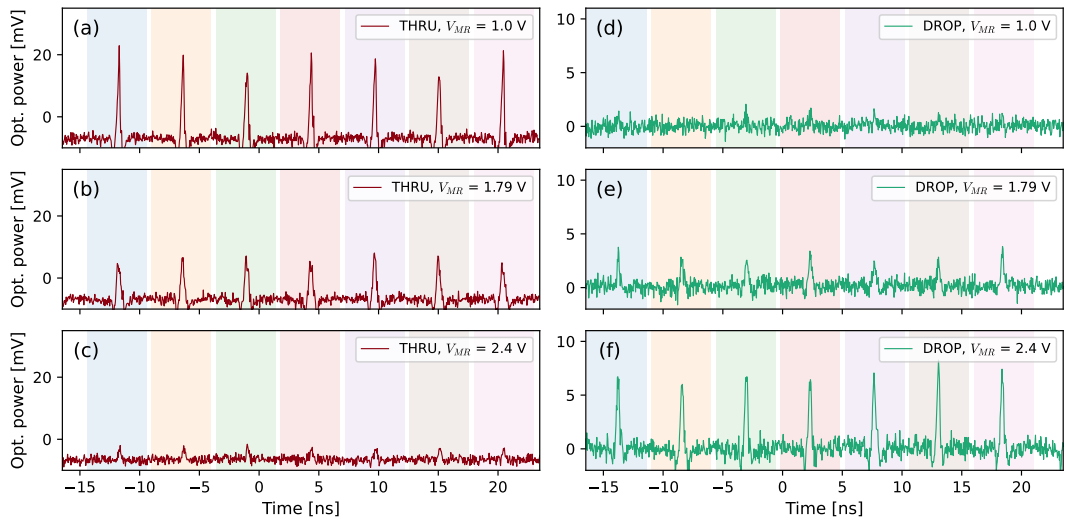


Figure 3.21: Sequences of deterministically triggered excitable spikes in the VCSEL-neuron after passing through the weighting MRR for three different values of V_{MR} . Traces were recorded on both outputs of the micro-ring: the THRU port (left column, in red) and the DROP port (right column, in green).

Fig. 3.21 shows temporal traces collected from the MRR THRU (in red) and DROP (in green) ports in the SiPh chip, when subject to an optical spiking signal (a total of $n_s = 7$ sub-ns long optical spikes with a repetition rate of 5 ns) from the VCSEL-neuron. The bias V_{MR} applied to the MRR heating element was swept from $V_{MR} = 1$ V to 3.9 V with 100 mV increments, total of $n_{\text{sweep}} = 29$ steps. The whole sweeping procedure was repeated a total of $n_m = 15$ times to provide ample data for better statistical evaluation of the experimental results.

During data post-processing, each of the $n_m = 15$ recorded traces (taken under

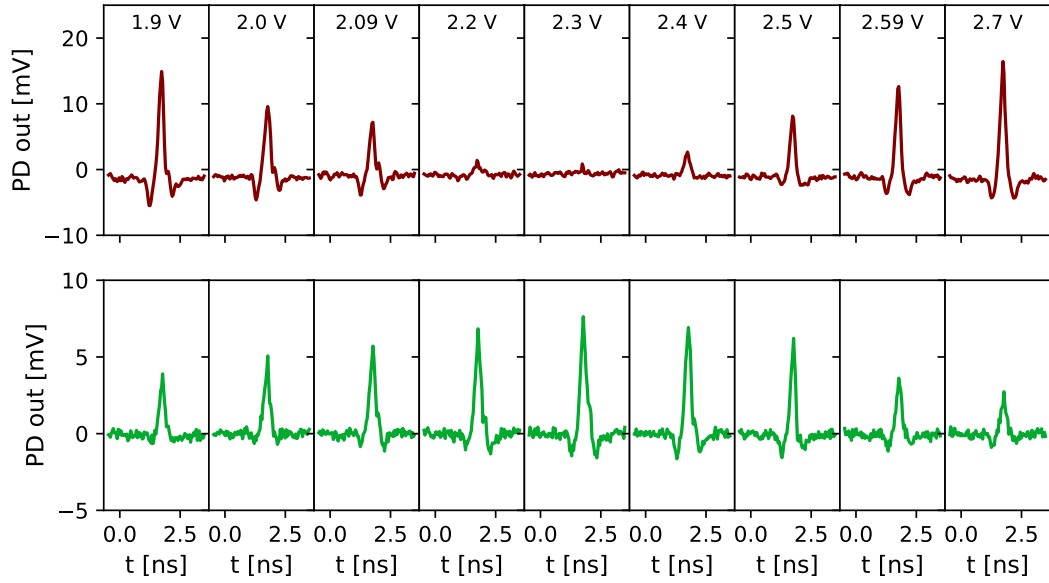


Figure 3.22: Figure of merit: spike amplitude modulation via MRR voltage tuning. (DC floor is removed).

equivalent experimental conditions) from both the THRU and DROP ports was sliced into small temporal windows (bins). This was done in such a way that each window contained a single MRR-weighted optical spike. These windows are shown as different colour shading in Fig. 3.21. Within each of these temporal bins, the spike maxima was located and utilised as alignment reference to calculate a representative, mean spike envelope for each DC bias value V_{MR} . This mean spike envelope from the THRU and DROP port is shown in Fig. 3.22 for nine different V_{MR} values applied to the MRR heater, ranging from 1.9 V to 2.7 V.

Furthermore, the peak amplitude of each representative mean optical spike from both the individual MRR output ports was calculated and plotted as a function of V_{MR} in Fig. 3.23(a,b). The latter also plots a Lorentzian fit for the interval around the extrema. The Lorentzian curve in Fig. 3.23(a,b) shows good fit to the experimental data. The signal recorded on the high-speed photodetector (Discovery LabBuddy) which was used to collect the THRU output port signal did not contain the DC signal component, representing the constant lasing (CW component) of the VCSEL-neuron operated above the lasing threshold. Therefore, this DC signal component was also numerically filtered

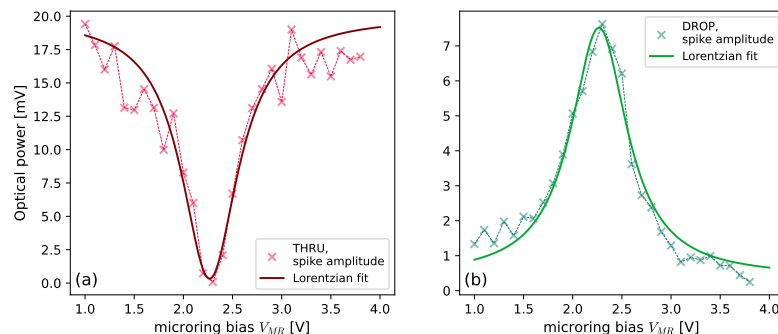


Figure 3.23: The figure of merit for spike weighting: spike amplitude modulation performed via MRR resonance tuning, as recorded from the THRU (in red) and DROP (in green) ports of the MRR. This tuning is performed by varying a bias V_{MRR} on a thermo-optic heater of the MRR. (The DC floor is removed during data analysis).

out from the photodetector response based on the DROP signal output, in order to provide comparison of signals of equivalent nature in both cases.

In conclusion, both the continuous single time traces shown in Fig. 3.21, the mean representative calculated optical spike envelopes in Fig. 3.22 and the Lorentzian-fitted mean optical spike amplitudes as a function of applied MRR ring heater bias V_{MR} in Fig. 3.23 demonstrate the viability of integrated photonic elements (used in optical weight banks) to perform synaptic weighting for high-speed optical neuromimetic signals generated by the all-optical photonic VCSEL-neuron.

3.5.2 Rate-coding in SiPh MRR-VCSEL-neuron system

In the second experimentally studied layout comprising the VCSEL-neuron and the SiPh MRR weight bank chip, the optical injection signal to enter the VCSEL-neuron is directly passed first through the MRR to control optical input strength into the VCSEL. In this approach, by setting the VCSEL-neuron at the edge of the locking/unlocking transition (where tonic spike firing can occur), it is possible to generate continuous spike trains. By modulating the injection signal power, the firing rate can be controlled. This, in turn, can be directly performed by the MRR in the integrated weight bank by tuning the voltage applied to the MRR heater V_{MR} . For more information on rate coding in VCSEL-neuron and utilisation of this phenomena for encoding of data, see Sec. 3.4.2

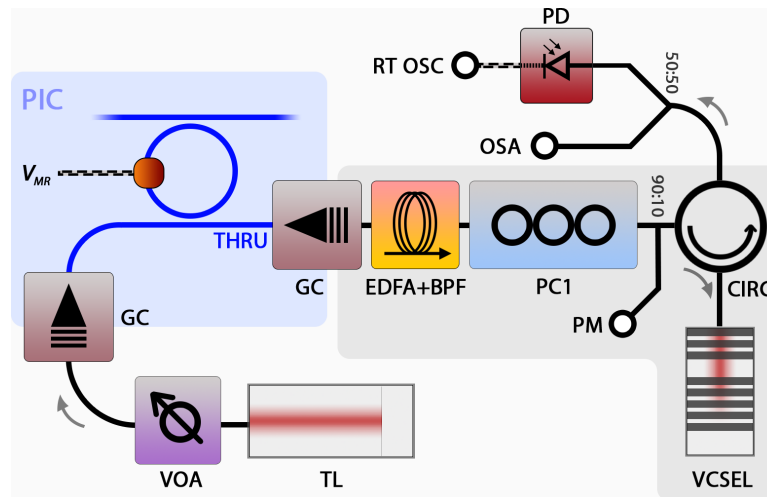


Figure 3.24: MRR \rightarrow VCSEL rate-coding experimental setup. TL - tuneable laser; VOA - variable optical attenuator; PC1 - polarisation controllers; PM - power meter; CIRC - circulator; VCSEL - vertical cavity surface emitting laser; EDFA - erbium-doped fibre amplifier; BPF - bandpass filter; RT OSC - real-time oscilloscope; OSA - optical spectrum analyzer; GC - grating coupler; PD - photodetector. Ratios for the power splitters/couplers are indicated in the figure.

and Sec. 3.4.3.

In this experiment, the SiPh MRR is used to controllably attenuate the injection signal (from an external, tuneable, CW laser) that locks into the VCSEL. To align the lasing emission spectrum of the VCSEL to the resonance of the MRR, the operation temperature of the VCSEL was adjusted via the thermistor built-in in the laser mount. Injection was then performed into the dominant polarisation mode of the VCSEL. Following the signal pathway, the optical output of a tuneable laser (EMCORE micro-ITLA) was passed through a variable optical attenuator (VOA) and coupled (through a vertical grating coupler) into the integrated circuit containing the weighting MRR device. The resonance frequency of the MRR was swept using a bias applied to the MRR adjacent heating element controlled by a programmable DC bias source. The THRU optical output of the MRR was then coupled off-chip (again through a vertical grating coupler) and collected into an optical fibre, amplified optically using an EDFA in series with a bandpass filter (BPF) to improve the signal-to-noise ratio (SNR) of the outgoing amplified optical signal. This signal was then optically injected into the VCSEL through an optical isolator. The optical injection signal was configured with an

initial frequency detuning of $\Delta\omega = -3.75$ GHz and an optical power of $P_{inj} \approx 220$ μ W. Finally, the output signal from the VCSEL-neuron was collected from the isolator and measured with a real-time oscilloscope using an amplified photodetector (Discovery Semiconductor LabBuddy).

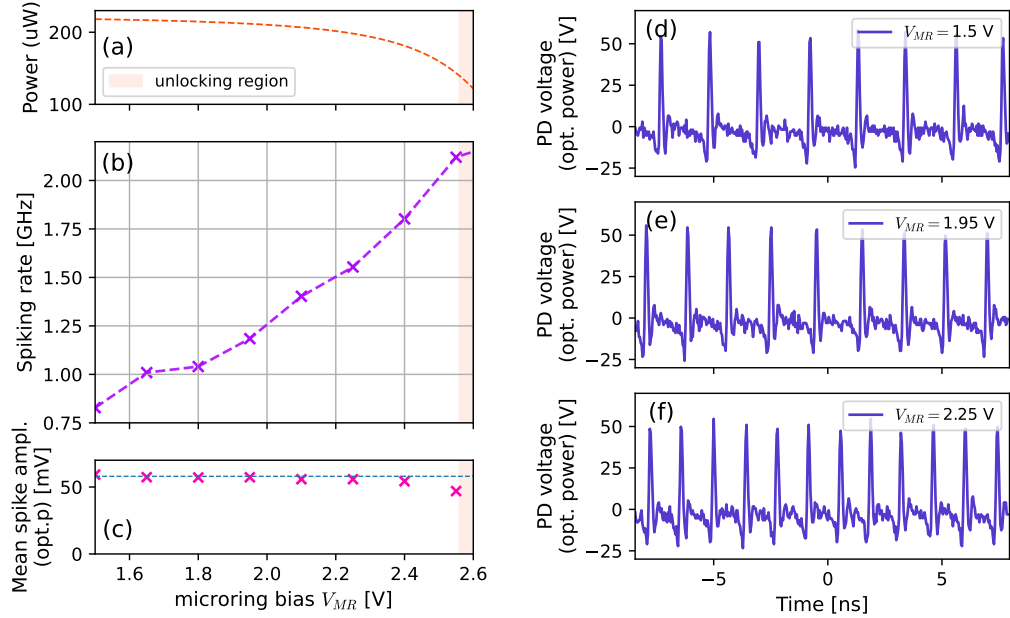


Figure 3.25: Demonstration of rate coding in a spiking VCSEL-neuron subject to injection of optical signals from a MRR. (a) Injection signal power as a function of V_{MR} . (b) Spike firing rate in the VCSEL-neuron as a function of V_{MR} . (c) Mean spike amplitudes within the optical spike train. (d-f) Examples of continuous spiking traces produced by the VCSEL-neuron locked to the signal from the MRR. As the voltage applied to the MRR heater is gradually increased from (a) 1.5 V through (b) 1.95 V to (c) 2.25 V, the spike firing rate monotonically increases.

The results from the MRR \rightarrow VCSEL experiment are shown in Fig. 3.25. Here, the DC voltage bias V_{MR} applied to the MRR heater was swept from 1.5 V to 4.2 V with 150 mV increments (total 21 steps). For each V_{MR} step, a 40 ns trace was recorded at the output of the VCSEL-neuron. In this trace, the occurring individual optical spikes were algorithmically counted to estimate the continuous firing rate for every value of applied MRR heater bias V_{MR} . This procedure was repeated $n = 12$ times and a mean optical spike firing rate as a function of applied V_{MR} was calculated from all the repetitions.

The optical power of the locking signal to the VCSEL as a function of V_{MR} is shown

in Fig. 3.25(a). As the MRR's resonance approaches the frequency of the optical signal injected into it, the power of the CW signal the output THRU port decreases. This CW signal is fed into the VCSEL as a locking signal. The VCSEL is operated under injection conditions such that it is near the locking/unlocking boundary. As the CW locking signal power from the THRU port decreases, the locked VCSEL's frequency moves closer to the locking boundary. This results in generation of a nonlinear dynamic response, which in this particular case corresponds to a change of spike firing rate for the constantly optically firing VCSEL-neuron [255]. If the locking signal power decreases too much, the injection locked VCSEL may completely unlock from the external injection, resulting in the disappearance of the spikes at the VCSEL-neuron's output.

As the optical power of the locking signal decreases, the continuous spiking rate at the VCSEL output increases. This is demonstrated in Fig. 3.25(b). At the start of the experiment, the VCSEL-neuron is locked to the MRR's (THRU) output signal (originally provided by a tuneable laser). It is also set into a continuous optical spike firing regime (further from the unlocking boundary). In such operational state, the continuous optical spike firing rate is approximately 750 MHz. As the ring resonance is approached during the V_{MR} sweep, the signal from the THRU port decreases in power, causing in turn an increase in the firing rate from the externally injection locked VCSEL-neuron to a maximum rate of ≈ 2.25 GHz (achieved at V_{MR} 2.55 V). For further decrease of injection power (obtained when the tuned ring resonance frequency closely approaches the frequency of tuneable laser signal), a transition is observed with a noticeable lowering of the oscillation amplitude (depicted in Fig. 3.25(c)), further followed by the VCSEL completely unlocking from the master signal (depicted as light orange shading in Fig. 3.25(a-c)). Therefore, only a reduced range of voltages is required for this type of functional layout to maintain the injection-locked state of the VCSEL-neuron. Examples of the rate-coded optical spike trains recorded from the VCSEL-neuron at various V_{MR} are shown in Fig. 3.25(d-f). In summary, the integrated silicon photonics MRR can be utilised to directly control the injection locking of a spiking VCSEL-neuron, therefore permitting its use as a high-speed, all-optical rate-encoding spiking system.

Chapter 4

Artificial spiking neurons based upon resonant tunnelling diodes (RTDs)

Relevant authored peer-reviewed manuscripts:

- ◆ Hejda, M., Malysheva, E., Owen-Newns, D., Al-Taai, Q.R.A., Zhang, W., Ortega-Piwonka, I., Javaloyes, J., Wasige, E., Dolores-Calzadilla, V., Figueiredo, J.M.L., Romeira, B. & Hurtado, A. *Artificial optoelectronic spiking neuron based on a resonant tunnelling diode coupled to a vertical cavity surface emitting laser*. Nanophotonics 0362. 10.1515/nanoph-2022-0362 (2022)

THIS chapter introduces the concept and properties of resonant tunnelling diodes (RTDs), a special type of active semiconductor device that utilises a quantum well within its heterostructure. As a class of mesoscopic devices, they allow us to directly utilise the phenomena of quantum tunnelling of electrons at room temperature, providing circuits that exhibit high-speed oscillations and dynamics. In this chapter, the RTD is experimentally and numerically studied as an excitable optoelectronic circuit, and the prospects of this system for neuromorphic information processing are investigated.

4.1 Overview, physics and band structure of RTDs

Resonant tunnelling diodes (RTDs), based upon a double-barrier quantum well (DBQW) semiconductor heterostructure, can be considered the simplest mesoscopic device for the demonstration of resonant quantum tunnelling phenomena. Whilst electrons in electronics are typically considered as corpuscular, their wave nature exhibits itself at sufficiently small scales (for which, the material lattice-induced scattering does not render the wavefunction incoherent). As an approximation for a semiconductor at room temperature, the electron phase-coherence distance is $\lesssim 10$ nm [266]. In RTDs, we observe quantum mechanical phenomena of electrons given that the thickness of the DBQW is below the phase-coherence length.

The typical features of a DBQW RTD band structure and its changes under the application of external forward bias can be seen in Fig. 4.1. The DBQW is defined by two very thin layers of undoped, wide-bandgap semiconductor material (such as AlAs [267]) sandwiching a narrow bandgap material (such as InGaAs [267]). The DBQW thickness is typically well below 10 nm and can be represented as a finite square quantum well with discrete, quantised electron energy levels. As long as the discrete well states are aligned (in resonance) with the source's conduction band, an electron (a quantum-mechanical wave) can tunnel through both the narrow energy barriers via the quantum well state. As the bias applied to the heterostructure increases (Fig. 4.1c), the bounded well state moves below the conduction band (within a few kT due to thermal fluctuations). In such situation, there is suddenly no available state for the electron to tunnel into, and the tunnelling probability (and therefore the current flowing through the RTD) suddenly exhibits a significant drop, appearing as a negative (decreasing) slope region within the I - V response curve of the device. This feature is typically referred to as NDR (negative differential resistance) region. The surrounding regions are commonly referred to as exhibiting positive differential resistance (PDR).

Thanks to the electron tunnelling effect and its transit times estimated at $\approx 10^{-13}$ s [268], RTDs exhibit a wideband, NDR region in their I - V characteristic persisting to room temperatures. While the common (positive) resistance consumes power

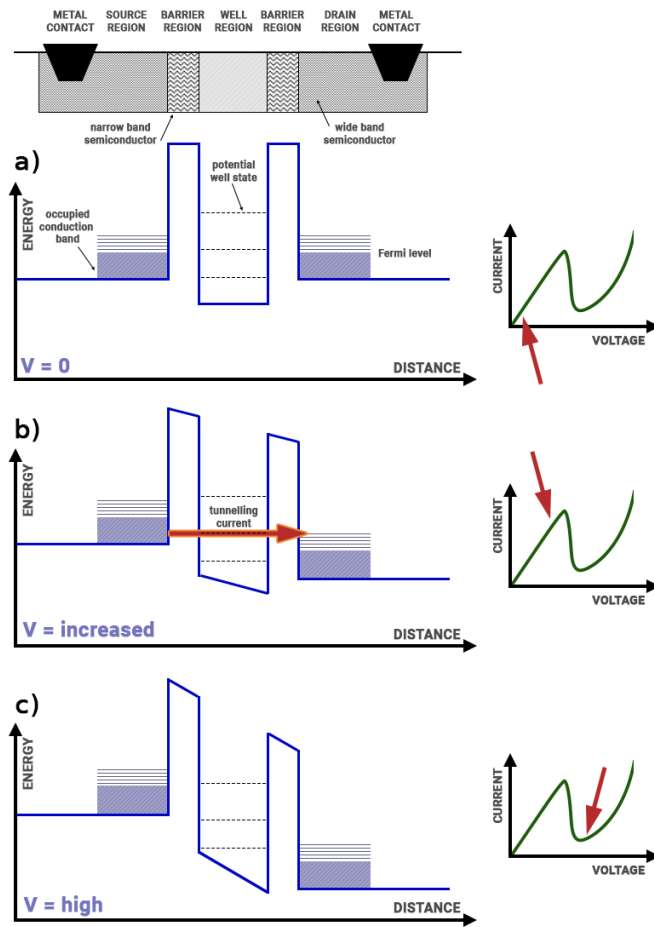


Figure 4.1: Simplified scheme of the resonant tunnelling diode (RTD) bandstructure, showing the double barrier quantum well (DBQW) sandwiched between two narrower-bandgap materials. a) No external bias is present. b) By applying an external voltage, the additional potential shifts the bands and aligns the resonant well state with the conduction band, allowing for resonant tunnelling through the DBQW, increasing the measured current. c) Once the applied voltage shifts the bands even further, the well state gets misaligned with respect to the conduction band, tunnelling probability drops to \approx zero and only the diode current is present, resulting in a drop of measured current in the I - V characteristic.

when current passes through it, a negative resistance generates power. Therefore, within the NDR region, the device exhibits electrical gain. As such, when the RTD is biased with a DC voltage V_0 within the NDR, electronic noise fluctuations are amplified. Therefore, with presence of electrical feedback (for example an inductor line, or other forms of parasitic components) and with sufficient gain to overcome losses, the RTD will oscillate [266]. Due to the very thin quantum well barriers and ultrafast nature of

quantum tunnelling, the RTD has a very high bandwidth and can produce oscillations that surpass THz frequencies. Notably, RTDs are also among the few quantum transport devices that can effectively operate at room temperatures [269].

Typically, planar RTD oscillators were based on an RTD incorporated within a slot metal resonator sustaining a standing electromagnetic wave and radiating RF power [270]. Spacers in the design serve to avoid unwanted dopant diffusion into the DBQW active region to lower the self-capacitance of the device while optimizing performance [271]. A self-synchronized optoelectronic oscillator with reduced phase noise based on an RTD-laser diode hybrid circuit with feedback was reported in [272].

Due to their intrinsic gain and photosensitivity, RTDs have been demonstrated as very sensitive photodetectors, down to single-photon signal levels [273]. RTD photodetectors with an AlGaAs/GaAs + GaInNAs absorption layer have been demonstrated to operate 1300 nm from picowatt power ranges [274] and a low-noise, single photon detector based upon an RTD with incorporated self-assembled InAs quantum dots has been demonstrated in [275, 276]. More detailed information on the ultra-high photodetection capabilities of RTDs can be found in a recent review [277]. Thanks to their ultra-high bandwidths, RTD structures are also suitable for operation in detection of THz radiation. An emitter-collector system of two RTDs operating at 300 GHz has been successfully demonstrated as a compact, reflection-based THz imaging system [278].

Going further towards information processing, branched RTD circuits have been demonstrated as universal logic gates [279]. A wireless communication system operating at 30-Gbit/s real-time error-free transmission without error correction has been demonstrated using an RTD oscillator with injection locking [280]. In general, RTD-based THz communication systems are among the simplest Tx and Rx circuit architectures in this frequency range [271].

Concerning their fabrication, RTDs are made using the methods of semiconductor growth and nanofabrication in III-V systems [281], with typical steps including resist coating, mask alignment and exposure, pattern development, metal deposition and lift-off, etching and passivation [282]. Some of the III-V material systems used for RTDs are reviewed in [271]. An example of the fabrication process for a simplified, THz-range

RTD oscillator circuit is shown in [283] with more details on the fabrication process available in [284].

This chapter summarises our results on the use of RTD devices as excitable functional elements for spike-based optoelectronic systems aimed towards neuromorphic information processing. All the devices in this thesis were fabricated by research collaborators at external institutions, namely the University of Glasgow in the United Kingdom [281] and Eindhoven University of Technology (TU/e) in the Netherlands. These institutions formed part of the consortium of the EU H2020 FET-Open ‘ChipAI’ programme that supported this PhD project.

4.2 Modelling (opto)electronic RTD circuits

In this section, the fundamental theory and RTD dynamical models are discussed. First, the electronic RTD dynamical model is introduced, followed by the extension into an optoelectronic architecture.

4.2.1 van der Pol & Liénard oscillators

A system that exhibits self-sustained oscillations can be understood as exhibiting a limit cycle (closed trajectory) around an unstable fixed point in the phase space. To describe a limit cycle, two variables are required [285] - and therefore, system motion within a limit cycle can be described by a 2D dynamical system (a second-order ordinary differential equation) [285]. A classical example of such system is the van der Pol oscillator. The van der Pol equation is a second-order, ordinary nonlinear differential equation with cubic nonlinearity devised by Balthazar van der Pol in 1927 to describe oscillations in electrical circuits containing triodes, coining the term “relaxation oscillations” for nonlinear oscillations produced by self-sustained oscillating systems [286]. The normalized van der Pol equation has a form of [285]:

$$\frac{d^2x}{dt^2} + \alpha(x^2 - 1)\frac{dx}{dt} + x + \beta y^3 = 0 \quad (4.1)$$

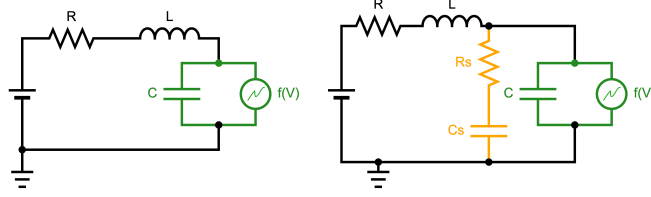


Figure 4.2: (left) Simplest RTD electronic circuit scheme. (right) Example of an RTD circuit with additional branch of shunting elements in orange. Equivalent elements of the RTD are highlighted in green. $f(V)$ denotes the nonlinear I - V .

where the parameter α governs the oscillation shape (harmonic to relaxation oscillations).

The simplest RTD equivalent lumped circuit is shown in Fig. 4.2. By utilising Kirchoff's laws on a two-terminal circuit [39], the circuit can be described using following differential equations:

$$\frac{dV(t)}{dt} = \frac{I(t) - f(V)}{C} \quad (4.2)$$

$$\frac{dI(t)}{dt} = \frac{V_0 - V_{RF} - RI - V(t)}{L} \quad (4.3)$$

where $V(t)$ is the voltage across the RTD and $I(t)$ is the total current. Indeed, by substituting a third-order polynomial into $f(V) = aV^3(t) - bV(t)$ into Eq. 4.2 and rearranging the terms accordingly, Eq. 4.1 is obtained. However, while the simple third-order polynomial for $f(V)$ does exhibit the key property of NDR, it does not fit very well the realistic I - V observed experimentally in DBQW RTDs (such as shown in Fig. 4.10). By using a more realistic, physics-based $f(V)$ as presented by Schulman [287] (described in greater detail in Sec. 4.2.2) and substituting, a generalized equation known as Liénard model is achieved [288].

$$\frac{d^2V}{dt^2} + h(V)\frac{dV}{dt} + g(V) = 0 \quad (4.4)$$

where $h(V) = \frac{R}{L} + \frac{1}{C} \cdot \frac{df(V)}{dV}$ is the damping factor, $g(V) = \frac{1}{LC} (V(t) - R \cdot f(V) - V_0 - V_{th})$ is the non-linear force and $f(V)$ is the non-linear I - V curve of the RTD. The van der Pol model is then a specific form of the general Liénard model for $h(x) = \alpha(x^2 - 1)$ and $g(x) = x$ [288]. Therefore, the set of Eqs. 4.3, 4.2 forms a Liénard oscillator model.

It's also notable that a Liénard oscillator subject to a time-dependent external force ($V_{ac} \neq 0$) provides an additional degree of freedom [289], resulting in a range of additional dynamical regimes besides self-sustained oscillations, including also excitability. The Liénard model also predicts a voltage-controlled oscillator (VCO) behaviour [288], where the oscillation frequency within the NDR depends on the applied voltage. In biology, the Liénard model has served as a model for part of nervous systems in crustaceans [290].

4.2.2 Nonlinear I - V characteristic of the RTDs

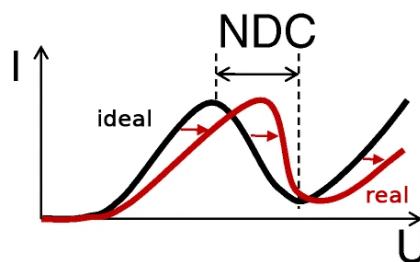


Figure 4.3: Comparison of an idealised RTD I - V (third-order polynomial, in black) with a more realistic I - V (in red). Reproduced and adapted from [291].

As was highlighted in the case of the van der Pol oscillator (Eq. 4.1), it was modelled based on arguably the simplest case of a nonlinear I - V characteristic: a third order polynomial [292] (lowest order exhibiting NDR). While such $f(V)$ provides the key property of NDR, it is a simplified approximation and does not fit the behaviour of a realistic RTDs particularly well. This is demonstrated in the schematic shown in Fig. 4.3. In general, electron transfer in the RTD system can be sufficiently approximated using two transfer channels for electrons: (1) the resonant transfer channel, and (2) the remainder of effects, such as thermal excitation over the barriers (all represented as a non-resonant channel) [266]. Following this separation into tunnelling-based current contribution and the remaining current-influencing effects, the Schulman [287] physical model of an RTD was introduced. In this model, the nonlinear I - V characteristic of a

DBQW RTD is determined using an analytical equation presented in [287], namely:

$$f(V) = A \ln \left(\frac{1 + e^{\frac{qe}{k_B T}(B-C+n_1V)}}{1 + e^{\frac{qe}{k_B T}(B-C-n_1V)}} \right) \times \left(\frac{\pi}{2} + \tan^{-1} \left(\frac{C - n_1V}{D} \right) \right) + H \left(e^{\frac{qe}{k_B T}n_2V} - 1 \right) \quad (4.5)$$

where T is the temperature, q is the electron charge and k_B is Boltzmann's constant; A , H are in A cm^{-2} (accounting for scaling using device area); B , C , D are in V; and n_1, n_2 are dimensionless. While these parameters enable us to fit the RTD I - V based on the physical system, the model parameter values unfortunately do not directly translate into heterostructure design parameters. The two terms describe how the resonant tunnelling effect defines the behaviour at lower voltages, while the classical diode character takes over at higher voltages [293].

Thanks to the highly nonlinear N -shape of the I - V characteristic, there is (at least one pair of) a local maxima and local minima. These points on the curve are commonly referred to as the I - V *peak* and the I - V *valley*, respectively. For operation in electronic circuits, high peak-to-valley current ratio (PVCR) is typically beneficial, since it leads to higher amplitude of non-linear dynamical responses. PVCR can be engineered by utilising specific quantum well barrier materials with high bandgap (high energy barrier) [266].

4.2.3 Influence of RTD circuit parameters

Eqs. 4.2, 4.3 can be further simplified by re-scaling using a factor dependent on the natural frequency of the RTD oscillator circuit. The resonant (tank) frequency of an LC circuit (the most favourable harmonic) can be calculated as:

$$\omega_0 = \frac{1}{\sqrt{LC}} \quad (4.6)$$

Therefore, together with a dimensionless time $t = \omega_0 \cdot t_{real}$, the rescaling factor μ is introduced as

$$\mu = \sqrt{\frac{C}{L}} \quad (4.7)$$

Using this factor, the dynamical equations therefore take a form of:

$$\mu \frac{dV(t)}{dt} = I(t) - f(V) \quad (4.8)$$

$$\mu^{-1} \frac{dI(t)}{dt} = V_0 - RI - V(t) \quad (4.9)$$

The scaling parameter μ is sometimes referred to as a *stiffness coefficient* and it governs the shape of the relaxation oscillations within the circuit [294]. This is of key importance for the excitable operation of RTD circuits, as very low μ ($\mu \ll R^{-1}$) gives rise to slow-fast dynamics. Slow-fast dynamics are observed when two or more variables of the dynamical system occur at very different time scales. More detailed explanation of this effect can be found, for example, in [294].

Capacitance

Capacitance C describes the capability of a component (circuit) to store energy in the form of electric charge. In RTDs, the intrinsic capacitance corresponds to: (1) geometric (emitter-collector) capacitance [289] due to charging/discharging of electrons in the depletion region and (2) capacitance of the quantum well [295]. Capacitance is dependent on the bias voltage due to charges stored in the quantum well [296]. In practical terms, the capacitance value in RTDs is mainly defined by the device structure and area. Low capacitance is desirable for high switching speed applications [269]. Expected values of capacitance per device area for regular RTDs are typically in the range of $C_s \approx 2 - 5 \frac{fF}{\mu m^2}$ [297].

Inductance

The inductance term L in the RTD circuit mainly comes from the circuit parasitics such as transmission lines (wires, cables [298], coplanar waveguides, microstrip lines [299]) in series with inductance from the lagging of quantum well charge [295]. The circuit inductance can be, to a certain extent, considered a free parameter as it not defined by the RTD epilayer structure, but rather by the oscillator design, geometry of transmission line and fabrication. Lower inductance is desirable, as it enables higher frequencies

in hybrid integrated RTD-LD circuits [288]. In the model, the refractory (lethargic) time between excitable spiking responses of the RTD decreases when inductance is lowered [289, 300].

Resistance

In the RTD circuit, R is a series resistance primarily related to the wires, cables [298] and the highly doped bulk regions around the DBQW [295]. Similarly to previous circuit parameters, minimised resistance R is desirable.

In addition, more complex circuit layouts can be utilised to optimize the circuit oscillation speeds. These include incorporation of shunting components such as resistors [301], capacitors [302] or combinations of both [298, 303] (example shown in Fig. 4.2), as well as use of using decoupling circuits [283].

4.2.4 Excitability in RTDs

As was highlighted in previous sections, the RTD is, as a generalized Liénard-type circuit, a dynamical system exhibiting a wide array of responses including harmonic and relaxation oscillations [22], mixed-mode-oscillations [304], stochastic resonance [279, 305] and most notably excitability [300].

The transition between harmonic and relaxation oscillation is, within the RTD model, governed by the stiffness coefficient μ . When μ is small ($\mu \ll 1$), the limit cycle consists of periods of buildups (slow phase) along the nullclines defined by the $f(V)$ function followed by quick discharge (fast phase). Therefore, while biased within the NDR region, the system exhibits slow-fast dynamics which are often referred to as relaxation oscillations (as the system quickly relaxes from the periods of slowly accumulated "stress" [22]). Furthermore, low μ systems can also exhibit excitability and spiking, which is a key dynamical behaviour utilised in this thesis for realisation of RTD-powered artificial optoelectronic spiking neurons. A parametric scan across two circuit parameters (C, L) is shown in Fig. 4.4. Each of the individual plots shows dynamical evolution of an RTD circuit that is biased in the valley operation point ($V_0 = 730$ mV) and being subject to a single input voltage perturbation (a 100 ps square

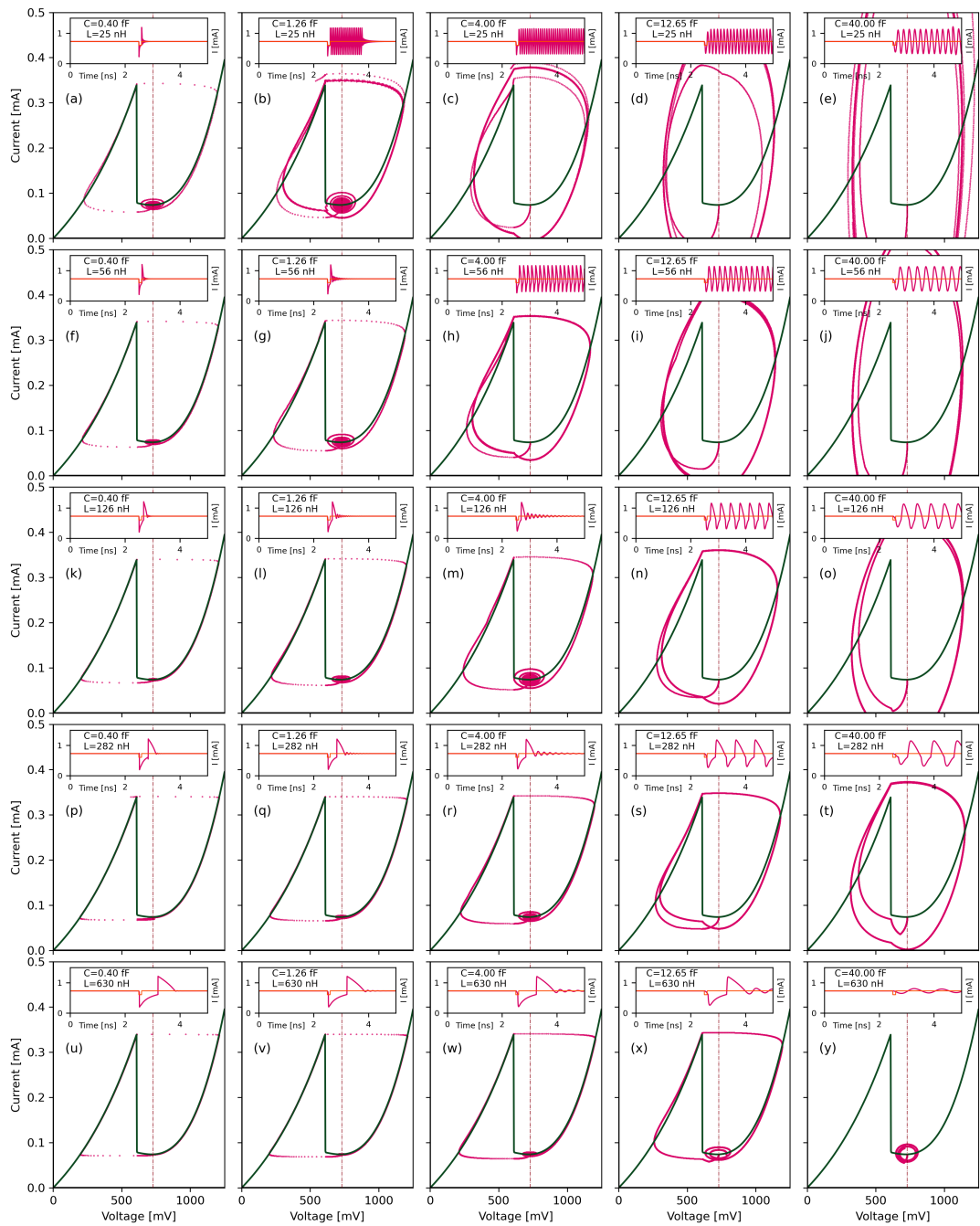


Figure 4.4: Dynamical evolution of an RTD circuit biased in the valley region of its I - V curve ($V_0 = 730$ mV) subject to a single voltage perturbation (a square pulse of 100 ps length and -125 mV amplitude) for variations of capacitance (along x -axis) and inductance (along y -axis).

pulse of -125 mV amplitude). In each case, a phase-plot (I - V) and a time trace (inset) is shown. Capacitance is varied left-to right across the plots, while inductance is varied top-to-bottom. The plots demonstrate that the fastest dynamics are observed for low values of both parameters. In the limit for $\mu \ll 1$, the system's spiking refractory time can be approximated as solely governed by the evolution along the slow limit cycle stages (adiabatic approximation), therefore being proportional to the inductance $t_{\text{ref}} \propto L$ [300]. Single spiking events are observed for (comparably) lower values of capacitance (such as in Fig. 4.4(f,k), among others), while continuous periodic oscillations are typically observed for higher values of C . A case of bursting is also observed in Fig. 4.4(b). For the highest investigated values of both C, L in Fig. 4.4(y), the input perturbation is not sufficiently strong to trigger any nonlinear dynamical response. This demonstrates how various dynamical regimes can be obtained in RTDs, and how circuit parameter engineering is crucial to enable excitable responses in these devices.

To deterministically trigger (spiking) responses in the RTD, a DC voltage is typically applied to the RTD such that the system is biased within close proximity of the PDR/NDR inflexion point, and then externally perturbed either with RF signals or optical pulses (in the case of photosensitive RTDs). Interestingly, as the N-shaped $f(V)$ exhibits two extrema (a local peak and a local valley), RTDs typically exhibit two different excitable regimes. Assuming a RTD circuit with an excitability-permitting set of R, L, C parameters, biasing the device in the PDR region within close proximity of the peak (first PDR/NDR inflexion) and perturbing towards the NDR yields one type of excitability. Biasing instead within the valley (NDR/PDR inflexion) in the 2nd PDR region and perturbing the system towards the NDR yields different excitable responses. In both cases, the perturbations must be of sufficient strength to exit the basin of attraction of the stable fixed points and reach the NDR region. This is typically referred to as "crossing the excitability threshold" and such perturbations are called super-threshold. Occurrence of excitability as well as its character (including its refractory time) are primarily governed by the shape (slope) of the I - V curve within each of the regions [300]. Since the limit cycle trajectory direction is fixed, spikes elicited in the peak and valley are mutually inverted in polarity. This dual-polarity, single-device

spiking capability of the RTD represents an interesting feature that is further explored in later sections of this thesis.

4.2.5 Extension to an optoelectronic model

In this thesis, RTD-based circuits are studied for the realisation of optoelectronic excitable systems. However, signals produced by these devices (in the minimal working circuit) are purely electronic (with the option for inherent light sensitivity in the device). Therefore, to utilise and operate RTDs with light-based input/output (I/O) signals (realising an Optical/Electronic/Optical system, O/E/O in short), the model and the studied system needs to be extended to account for opto-electronic (O/E) conversion at the input, and electronic-optical (E/O) conversion at the output. There are multiple different approaches to such model extension, and options for both inputs and outputs will be discussed at a later stage in this thesis. Particularly, the following sections will describe various approaches to O/E and E/O conversion in the RTD systems. The O/E approaches include photo-detecting RTD nodes (pRTD) and systems with RTD elements coupled to photodetectors (RTD-PD nodes), whilst the latter include RTD elements coupled to light emitting devices, such as laser diodes (LD), to form RTD-LD nodes.

pRTD devices & RTD-PD nodes

First, let's consider the case for optical input into RTD circuits. The RTD semiconductor layerstack can, when appropriately engineered, sense light directly within the heterostructure. This represents an interesting alternative to other more common photosensitive devices thanks to the RTD's intrinsic gain, low voltage operation, high responsivity (down to single photon levels [277]) and high gain-bandwidth product [306]. A pulse of light with photons of sufficient energy incident on the RTD structure (embedded with sufficiently thick light absorbing layer) produces photo-induced excitons, and under the influence of an external bias-induced electric field, the holes accumulate at the DBQW barrier while the electrons are gathered at the collector (external circuit) [307]. These accumulated charges negate the field within the absorbing layer and locally enhance the

field within the DBQW, resulting in a shift of the I - V curve towards lower voltages [306]. The effects of direct illumination on a RTD and its photoconductivity and charge accumulation effects on the I - V characteristic were studied in the work of Coelho et al. [308]. In this work, the photoinduced voltage (together with the equivalent series resistance) directly influences the analytical RTD equation [287] by incorporating a photo-induced I - V shift to lower voltages. Photosensitive RTD oscillators circuits can be operated in multiple different modes based on the bias point [277].

$$C \frac{dV}{dt} = I - f(V) - I_{ph} \quad (4.10)$$

By assuming a static $f(V)$ without light-induced changes and a low input optical power level (with small power variations), we can utilise a simplified, linearized sensitivity-power relation in the PD term [274]. Therefore, the photocurrent term I_{ph} in Eq. 4.10 can be considered as:

$$I_{ph} = \kappa S_m(t) \quad (4.11)$$

RTD-LD nodes

Previous works have shown experimentally the coupling of RTD elements to laser diodes (LDs) to form RTD-LD hybrid circuits. For example, coupling a multimode DFB-type laser with an RTD device was demonstrated, with the RTD and LD wire bonded together and connected in series using a printed circuit board (PCB) [272, 288].

In this thesis, a simple laser diode (LD) model is considered without the incorporation of polarisation dynamics. This model is derived from the standard rate model describing the dynamics of photon number $S(t)$ and carrier number $N(t)$ [309–312]. This laser model assumes linear gain $G = \gamma_m(N - N_0)$ and that the LD is directly coupled to the RTD (in series) via the current term I :

$$\frac{dS}{dt} = \left(\gamma_m(N - N_0) - \frac{1}{\tau_p} \right) S + \gamma_m N + \sqrt{\gamma_m N S} \xi(t) \quad (4.12)$$

$$\frac{dN}{dt} = \frac{J + \eta I}{q_e} - (\gamma_l + \gamma_m + \gamma_{nr})N - \gamma_m(N - N_0)|E|^2 \quad (4.13)$$

Here, q_e is the electron charge. N_0 is the transparency carrier number, τ_p is the photon lifetime (≈ 500 fs in a realistic nanolaser low-Q cavity designs [313]), γ_m is the spontaneous emission rate into the lasing mode (where $\gamma_m \cdot S$ is the stimulated emission rate), γ_l is the radiative decay rate into the leaky modes, γ_{nr} is the non-radiative spontaneous emission coefficient. In our modelling, we assume a low-footprint nanolaser design, which represents a main goal of the EU 'ChipAI' project supporting my PhD project. However, the model parameters can also be adapted to describe other more conventional semiconductor lasers, such as Vertical-VCSELs, DFB lasers or Fabry-Perot lasers, among others. This model assumes a case of independent, additional biasing for the laser diode (via the J term) in addition to the biasing provided to the RTD. In some models, such as the monolithic, single nanopillar implementation, such decoupled control may not be feasible as the nanolaser current will directly correspond to the RTD bias current level. The stochastic nature of the system is represented by the multiplicative noise term $\sqrt{\gamma_m N S} \xi(t)$, where $\xi(t)$ is a time-uncorrelated white noise function.

RTD-LED nodes

It is worth mentioning that as an alternative to the RTD-LD arrangement, a light-emitting diode (LED) may also represent a viable, less fabrication demanding light source for operation in conjunction with an RTD towards optoelectronic spiking sources, particularly in cases that do not require coherence or long distance signalling. In some cases, nanoLEDs can show even better performance in terms of noise reduction compared to nanolasers [314]. The prospects of spiking, nanoscale (sub- λ) monolithic RTD-LEDs were thoroughly analyzed in [150], with numerical results demonstrating multi-GHz spike rates activated upon receiving exceptionally low (sub-10 mV) electronic perturbation signals. Whilst nanoLEDs may reach a fundamental energy/speed limit for data rates exceeding a few gigahertz, sub- λ nanolasers (with active dimensions in the range of few hundred nanometers) may enable direct modulation rates higher than 40 Gb/s. Therefore, such systems would satisfy the two key operational metrics of prospective RTD-based neuromorphic optoelectronic systems: high operational speeds and low power requirements. It is worth noting that in the scope of integrated photonics

information processing systems that would utilise RTD-based optoelectronic nodes, a combination of more robust LED-based nodes (for intra-chip signalling) and laser-based nodes (for longer distance and/or off-chip signalling) is a viable potential solution.

4.3 Numerical analysis of a monolithic O/E RTD (PRL-MN)

The numerical model of an O/E/O RTD circuit is realised by combining the rescaled Lié-nard oscillator-based dynamical equations (Eqs. 4.8, 4.9 that incorporate the simplified photodetection term (Eq. 4.11) together with the laser diode equations (Eqs. 4.12, 4.13). This model assumes a serial connection between the RTD and the optoelectronic elements, and uses experimentally-measured I - V that is representative of a nanoscale RTD. Similarly, the laser model parameters are representative of a nanoscale laser. This model therefore represents a monolithic, nanoscale RTD device that incorporates the photosensitive semiconductor layer stack of the DBQW alongside the laser cavity, with the DBQW being at the top of a monolithic nanopillar structure. This arrangement will be referred to as a PD-RTD-LD monolithic node (PRL-MN).

By combining Eqs. 4.8, 4.9 with the photodetection term (Eq. 4.11) and the laser diode equations (Eqs. 4.12, 4.13), we obtain:

$$\mu \frac{dV(t)}{dt} = I(t) - f(V) - \kappa S_m(t) \quad (4.14)$$

$$\mu^{-1} \frac{dI(t)}{dt} = V_0 - RI - V(t) \quad (4.15)$$

$$\frac{dS}{dt} = \left(\gamma_m(N - N_0) - \frac{1}{\tau_p} \right) S + \gamma_m N + \sqrt{\gamma_m N S} \xi(t) \quad (4.16)$$

$$\frac{dN}{dt} = \frac{J + \eta I}{q_e} - (\gamma_l + \gamma_m + \gamma_{nr})N - \gamma_m(N - N_0)|E|^2 \quad (4.17)$$

where V is the voltage along the RTD, $I(t)$ is the total current in the circuit, $S(t)$ is the photon number and $N(t)$ is the carrier number.

Alternatively, the laser (Eqns. 4.16, 4.17) within the model can be coupled into

the RTD (Eqns. 4.14, 4.15) through the voltage term instead of the current term. In this arrangement, the model's dynamical response mimics more closely the dynamical behaviour observed in the PRL and nano-pRL experiments, where the RF signals from the RTD are AC-coupled directly to the laser through a bias-tee network. This is done by modifying Eq. 4.17 as:

$$\frac{dN}{dt} = \frac{\eta V}{q_e R_0} - (\gamma_l + \gamma_m + \gamma_{nr})N - \gamma_m(N - N_0)S \quad (4.18)$$

The set of ordinary differential equations (ODEs) is solved numerically, using time-varying signals for input variables (electrical input $V(t)$, optical input $S_m(t)$). The I , V equations can be solved using an explicit Runge-Kutta 5th/4th mixed order method (available in Python for example in the form of an initial value problem (IVP) solver `scipy.integrate.solve_ivp` [315]), while the complete set of equations in this work was solved using a custom-implemented symplectic Euler method. The numerical implementation of the symplectic Euler method was originally provided in Matlab by Prof Julien Javaloyes from the University of Balearic Islands (UIB), Spain. This code was then converted to object-based Python syntax and accelerated using a high performance just-in-time (JIT) compiler [316].

4.3.1 Parameters of the PRL-MN model

The complete set of parameters used within this numerical model can be divided into subgroups governing particular parts of the model.

The I - V parameters

Following the physical model introduced by Schulman [287], there are 7 parameters governing the shape of the I - V characteristic. The Equation for this model has been previously introduced in Sec. 4.2.2.

This set of parameters is based on a physical, nanoscale RTD. This set of parameters yields a highly nonlinear I - V curve that is shown in Fig. 4.5. This curve exhibits a single region of NDR sandwiched between two PDR regions, with boundaries at 608 mV

Table 4.1: Modelled RTD I - V curve parameters

Parameter	Value	Units
A	-5.5×10^{-5}	A
B	0.033	V
C	0.113	V
D	-2.8×10^{-6}	V
N_1	0.185	-
N_2	0.045	-
H	18×10^{-5}	A

(peak) and 719 mV (valley).

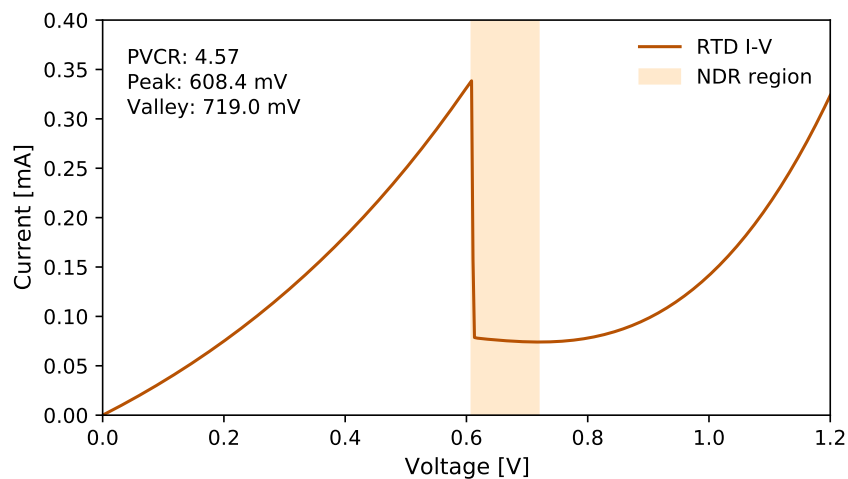


Figure 4.5: Current-voltage (I - V) characteristic of the modelled RTD device with the NDR highlighted (in orange).

To provide better understanding of the effects of the parameters in the Schulman model, Fig. 4.6 provides a visualisation of how variations in different $f(V)$ parameters influence the curve shape.

The circuit parameters

The RTD forms part of an oscillator electronic circuit (Fig. 4.2), as discussed in Sec. 4.2.3. The electronic circuit parameters have significant influence over the non-linear dynamics of the circuit, as demonstrated in Fig. 4.4. The circuit is considered as a small-signal equivalent circuit with 7 parameters that can be considered as influencing the operation

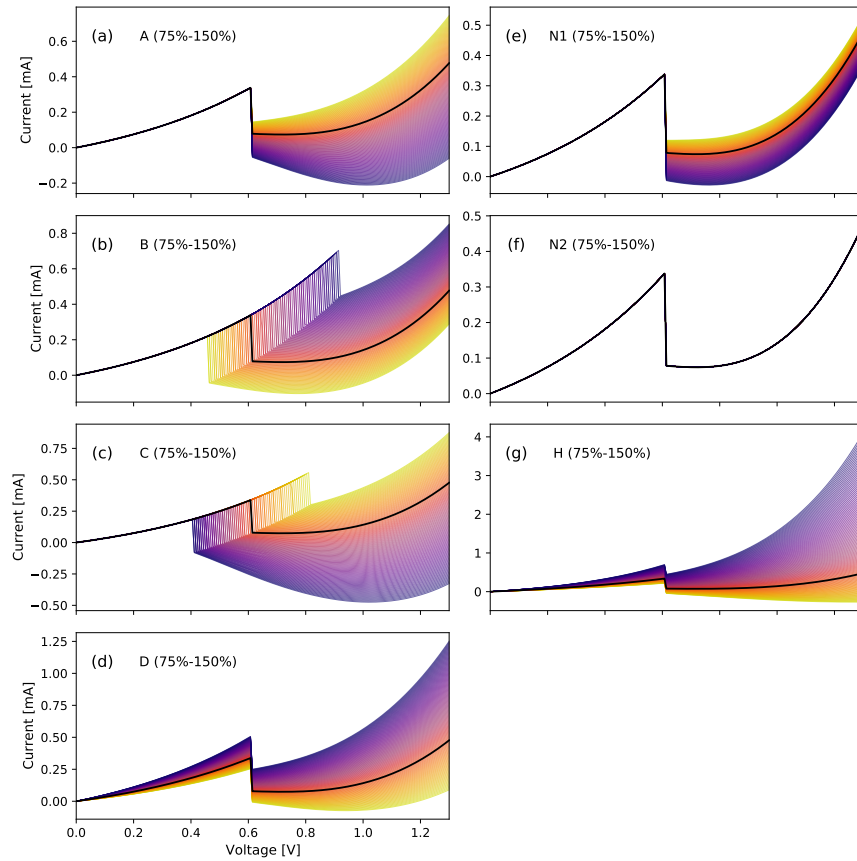


Figure 4.6: Scanning I-V curves of an RTD. The starting values correspond to a nanoscale RTD device, and parameter value scans are performed between 75% and 150% of starting values. In the case shown in (f), no significant variation is observed for N2 under the current set of parameters.

of the circuit: DC bias voltage V_0 , a series resistance R , inductance L and intrinsic capacitance C . κ represents a simplified conversion factor for optical-to-electrical signals. Typically, this accounts for the photodetection process responsivity (sensitivity), representing the rate of photocurrent generation per photon count (intensity). In the model, we represent responsivity as a constant value. Finally η represents the efficiency of current coupling from the circuit into the laser.

Laser model parameters

For fully optoelectronic operation of the RTD-powered circuit, a light source is coupled to the model. The laser diode is incorporated via a standard rate model with linear

Table 4.2: Modelled RTD circuit parameters

Parameter	Value	Units
V_0	<i>adjustable</i>	V
R	10	Ω
R_0^\dagger	50	Ω
L	126×10^{-9}	H
C	2×10^{-15}	F
κ	4×10^{-7}	A
η	0.2	-

[†]For V-coupled laser model

gain and multiplicative noise, describing photon number S (calculated from complex electric field $S = |E|^2$ [313]) and carrier population N .

In total, six parameters are related to the laser model. N_t is the transparency carrier number. α is the linewidth enhancement factor that couples intensity and phase fluctuations via the carrier dependent refraction index. τ_p is the photon lifetime. The parameter γ_m corresponds to the spontaneous emission rate into the lasing mode while γ_l represents the radiative decay rate into the leaky modes and γ_{nr} is the non-radiative spontaneous emission coefficient. The final parameter J represents an additional current term provided to the laser alongside the current from the RTD. This is used to adjust the steady-state operation point of the laser, and can be used for output spike amplitude modulation.

Table 4.3: Modelled RTD-LD circuit: laser parameters

Parameter	Value	Units
N_t	5×10^5	-
α	2	-
τ_p	5×10^{-13}	s
γ_m	1×10^7	s^{-1}
γ_l	1×10^9	s^{-1}
γ_{nr}	2×10^9	s^{-1}
J	3×10^{-4} (<i>variable</i>)	A

The spontaneous emission factor $\frac{\gamma_m}{\gamma_m + \gamma_l} = 0.0099$ is a value representative of a low-Q

cavity laser at the nanoscale, enclosed in a metallic cavity. It is worth mentioning that the experimental RTDs (Secs. 4.4, 4.5) rely on coupling to a VCSEL, which represents the case of DBR cavity with a multiple quantum well (MQW) active medium. Meanwhile, this nanoscale laser model represents a device with bulk active material and metal-dielectric cavity. Despite their different principles of operation, the lasers in both the experimental and numerical case represents an O/E transmitter, a functionality that is satisfied by either of the devices.

4.3.2 Spiking and excitable threshold in the PRL-MN model

First, a clearly defined excitability threshold and spiking response is demonstrated in the PRL-MN. The response of the model is shown in Fig. 4.7. In this case, the RTD is operated in the valley close to the inflexion point ($V_0 = 730$ mV). Fig. 4.7(a) demonstrates the modulation signals applied to the RTD, both in the electronic domain as an RF signal (in green, here no modulation is present), and optical perturbations (in pink). The optical perturbations have a form of 60 ps square pulses that gradually increase in amplitude ($S_1 = 75$, $S_2 = 150$, $S_3 = 225$, $S_4 = 300$). Fig. 4.7(b-e) shows the response of the circuit in terms of voltage, current, photon number and carrier number respectively. Since the RTD in the model is operated in valley point with a shallow character, we observe more pronounced relaxation oscillations when the system is perturbed. However, only a single (strongest) input optical perturbation produces a full-fledged excitable spiking response in both voltage and current.

Fig. 4.8 provides more detailed insight into the dynamics happening in the valley for positive input optical perturbations. These phase plots demonstrate how the dynamical system reacts to each of the four gradually increasing perturbation pulses from Fig. 4.7(a). The phase portraits reveal that for small (subthreshold) perturbations near the stable fixed point, the resulting trajectories do not reach the NDR boundary and instead relax back to the steady state via dampened harmonic oscillations. For a super-threshold perturbation (Fig. 4.8(d)), the system is perturbed sufficiently to reach the excitability boundary in the phase space, which here is approximately equivalent to the steep NDR section. Once the system reaches this point, a full excitable response following a large

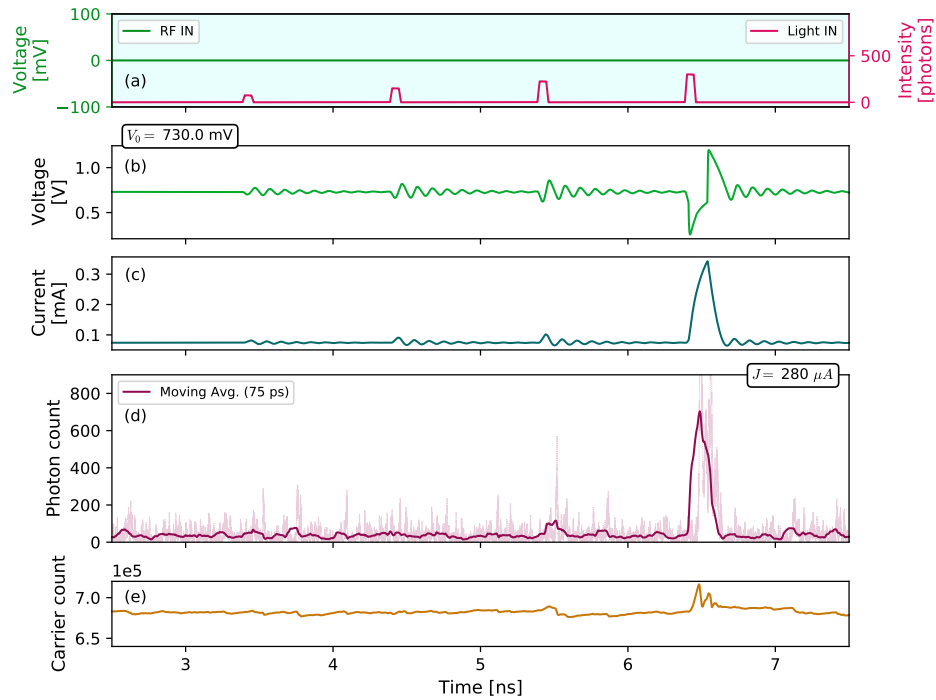


Figure 4.7: Thresholding demonstration in the PRL-MN numerical model, with traces (top to bottom) demonstrating input perturbations (red for optical, green for electrical), voltage, current, photon count and carrier number respectively.

phase-space excursion with slow-fast regions is triggered. This point is denoted with an asterisk in Fig. 4.8(d).

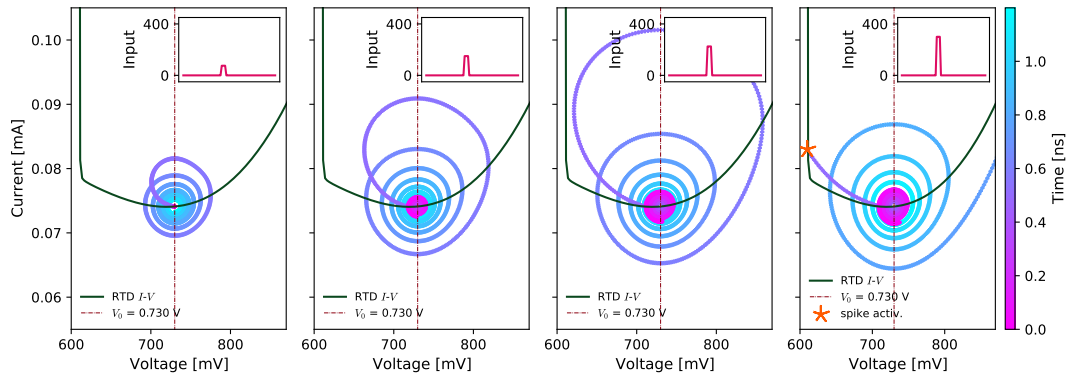


Figure 4.8: Phase space plots for gradually increasing optical perturbations in the valley operation point of the RTD. The asterisk in the right-most figure represents the moment when a full excitatory spike is elicited in response to a sufficiently strong perturbation.

4.3.3 Pre-synaptic spike weighting in the PRL-MN model

Thanks to the adjustable additional bias current J applied to the laser diode, the operational point (pump current) of the laser element within the model can be controlled. This allows for direct tuning of the output optical spike amplitude (solely by varying the J term), effectively realising a pre-synaptic, on-demand spike amplitude weighting functionality.

First, a spike has to be elicited in the PRL-MN node. To achieve this, the valley-biased RTD ($V_0 = 730$ mV) is optically perturbed using a single, 100 ps long square perturbation, responding with an excitable spike that is directly current coupled to the nanolaser element. The approximate threshold current for the utilised set of laser model parameters [313] is $I_{thr} \approx 337$ μ A. To perform this demonstration, the J value is varied from 200 μ A (far sub-threshold) to 300 μ A (close to threshold) with 10 μ A increments. The results for this demonstration are shown in Fig. 4.9. The time traces showing the optical output of the PRL-MN for variations in J are shown in Fig. 4.9(a-j). These are shown as 75 ps simple moving averages to filter the ultra high-frequency variations in the physical model. It can be clearly seen that variations in the J term result in a range of different optical spike amplitudes at the output of the PRL-MN. As the model is biased closer towards the threshold and the operation moves further from the gain-switching-like operation, the noise floor becomes more pronounced in the signal. Fig. 4.9(k) shows the mean captured spike amplitude taken over $n = 20$ repetitions. Once the laser is operated close to or above threshold, the spike amplitude with respect to the signal floor remains fixed. This is demonstrated by the saturation-like effect on the depicted trend at $J \approx 270$ μ A (this value is below I_{thr} due to RTD current contribution). Finally, a signal-to-noise ratio (SNR) plot is provided, using the formula

$$\text{SNR} = 20 \log_{10} \frac{\overline{S}_{\text{spike}}}{\overline{S}_{\text{floor}}} \quad (4.19)$$

for each value of J , showing maximum SNR further below threshold at $J = 240$ μ A.

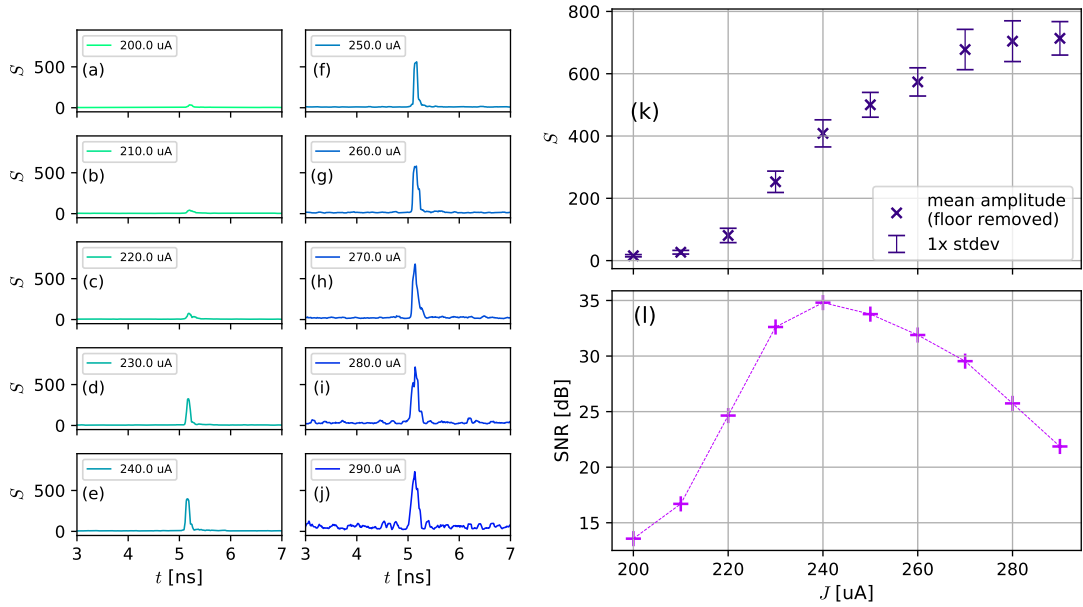


Figure 4.9: Demonstration of spike weighting in the numerical RTD-MN model. (a-j) Time traces of PRL-MN optical output, showing single optical spike for variations in J . (k) Mean spike amplitudes (signal floor subtracted) obtained over $n = 20$ repetitions. (l) Signal-to-noise (SNR) for obtained spikes.

4.4 Experimental analysis of μ -RTDs (PRL nodes) for spiking neuronal models

The remainder of this chapter will present experimental measurements of RTD-powered optoelectronic circuits, with primary focus on excitable spiking and related dynamical phenomena. Within the scope of this chapter, two different terms are used that refer to two different RTD-powered architectures:

- PD-RTD(-LD) (or PRL in short), an electronic circuit with a μ -scale RTD coupled serially to an active, amplified photodetector (PD8GS, Thorlabs) and (optionally) a RF-linked telecom-wavelength VCSEL. The PRL circuit is experimentally studied in this Sec. 4.4.
- pRTD(-LD) (nano-pRL in short), an electronic circuit with a nano-scale RTD with embedded optical window and a RF-linked telecom-wavelength VCSEL

(optional) that can directly respond to externally-injected optical input signals.

The nano-pRL circuit is experimentally studied in Sec. 4.5.

In both of these cases, the architectures were realised using an RTD coupled to off-the-shelf optoelectronic and fibre-optic components. These architectures can be considered modular (e.g. the VCSEL can optionally be disconnected in the layout).

4.4.1 Device characteristics, experimental setup: PD-RTD-VCSEL (PRL)

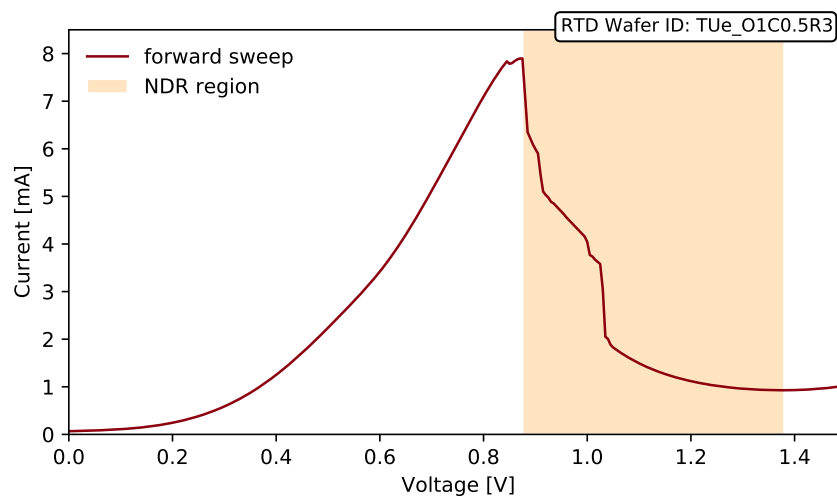


Figure 4.10: Current-voltage (I - V) characteristic of the used micro-RTD device (forward biased) with the NDR region highlighted (in orange).

All the experimental analyses of the PD-RTD-LD (PRL) layout were realised using a micro-RTD device fabricated by our 'ChipAI' EU project partners at the Eindhoven University of Technology (TUE) in the Netherlands (E. Malysheva and V. Dolores-Calzadilla). These RTDs are fabricated using molecular beam epitaxy (MBE) on a semi-insulating InP substrate incorporating a DBQW in its structure. The AlAs barriers are 1.7 nm thick, and the InGaAs well is 5.7 nm thick. The DBQW is surrounded by highly doped n -InGaAs contact layers. The full layerstack can be seen in Table 4.4. A micrograph of the wafer with individual RTD devices is shown in Fig. 4.11

The characterisation of the electrical properties of the μ -RTD reveals a highly

Table 4.4: Epilayer stack for the μ -RTD device. This device was designed and fabricated by team of our project collaborators at the Eindhoven University of Technology (E. Malysheva, V. Calzadilla).

Layer	Thickness (nm)	Semiconductor	Doping (cm^{-3})	Description
12	200	InP	n.i.d	Cap layer
11	100	n-In _{0.532} Ga _{0.438} As	$N = 2e19$	N contact (emitter)
10	50	n-InGaAs	$N = 5e16$	N- doped layer
9	2	InGaAs	n.i.d	Dopant diffusion stop layer
8	1.7	AlAs	n.i.d	DBQW barrier
7	5.7	InGaAs	n.i.d	Well
6	1.7	AlAs	n.i.d	DBQW barrier
5	2	InGaAs	n.i.d	Dopant diffusion stop layer
4	50	n-InGaAs	$N = 5e16$	N- doped layer
3	50	n-InGaAs	$N = 5e18$	N+ doped layer
2	20	n-InP	$N = 5e18$	N doped layer (etch stop)
1	100	n-InGaAs	$N = 1e19$	Collector

nonlinear I-V characteristic (shown in Fig. 4.10) with a pronounced NDR region starting around 900 mV and extending beyond 1.25 V as highlighted with orange shading in Fig. 4.10. The RTD exhibits a large peak-to-valley current ratio (PVCR ≈ 8.5). High PVCR is desirable as it enables dynamical responses (such as spiking) with high amplitude contrast.

The wafer containing the micron-scale RTDs was probed using a custom-built radio frequency (RF) probing and characterisation station, utilising precision micropositioners (Signatone S-M40) and RF probes (GGB Picoprobe Model 40A GSG). The RF probe was connected using a short interconnect cable (Mini-Circuits Hand-Flex) to the RC+DF port of a 12 GHz bias-tee (Inmet 8800SMF1-12). The bias tee was used to provide the required DC bias (V_0) that sets the desired biasing operation point of the RTD (near the peak in the 1st PDR, or near the valley in the 2nd PDR). The RF output was connected using an SMA cable to a resistive power splitter that allows us to combine together the inputs and outputs of the probed RTD device.

The optical input functionality to the RTD is realised with a 9 GHz amplified InGaAs

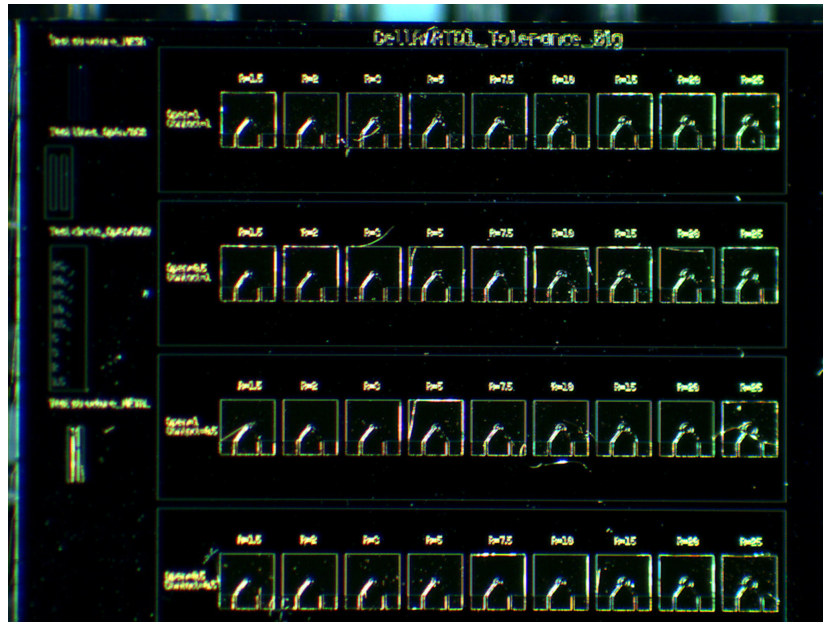


Figure 4.11: A microscope image of the wafer with μ -RTD devices used in the experiments. These were fabricated by E. Malysheva (V. Dolores-Calzadilla's group) at the Eindhoven University of Technology (TUE).

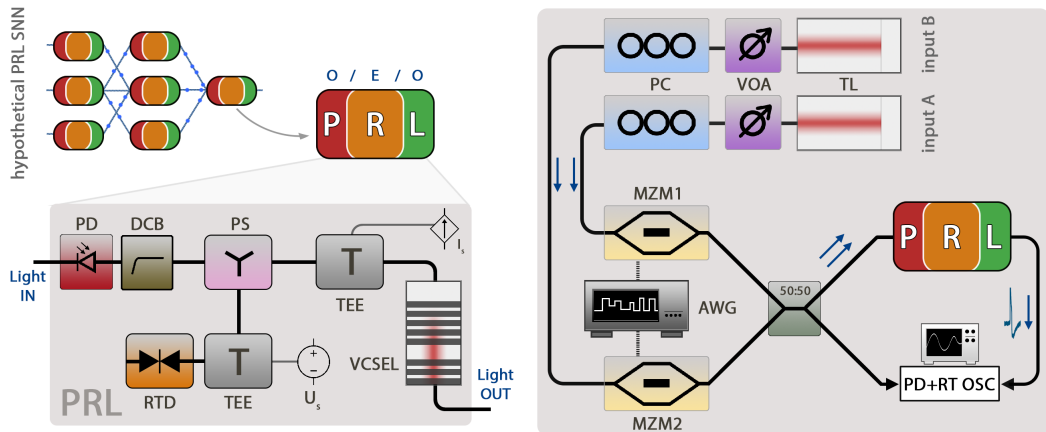


Figure 4.12: Experimental setup for the characterisation of the opto-electronic dynamical spiking regimes in the module PRL architecture.

photodetector (PD, Thorlabs PDA8GS). The PD was used to detect the incoming input optical signals and convert them to the required RF modulating signals entering the RTD for deterministic elicitation of spikes. Thanks to the amplified photodetector used in this layout, there is no need for additional amplification (as was the case in previous works [317]). The photodetector is coupled to one of the splitter ports through

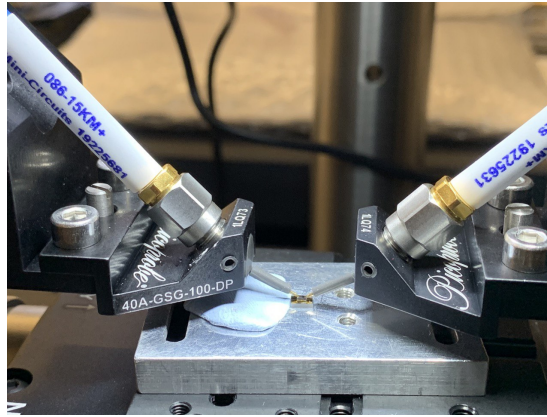


Figure 4.13: Photograph of the probe station used to drive RTD circuits. Here, a two-terminal device is probed using RF probes attached to micro-positioning systems.

a DC blocking element (DCB in Fig. 4.12a) to filter out any constant amplitude optical signals. Therefore, the PRL node only reacts to varying signals (such as input pulse or spikes) and the RTD operational point V_0 is prevented from unwanted shifting due to such signals.

The optical output of the PRL node is realised using an electrically modulated VCSEL. The VCSEL is a commercially-sourced, telecom device operating at the wavelength of 1550 nm. The VCSEL was driven in a laser mount (Thorlabs LDM56) with thermal control via a thermistor ($R = 12.5 \text{ k}\Omega$, corresponding to 293 K) and with a 50Ω RF input that is AC-coupled directly to the VCSEL through a bias tee (up to 500 MHz). The lasing threshold of the VCSEL was measured to be approx. 1.9 mA at room temperature ($T = 298 \text{ K}$), and the laser was operated with current of 2.2 mA, slightly above the lasing threshold. The readout from the VCSEL was recorded through an optical isolator (to avoid unwanted reflections and increase the S/N ratio) on a 16 GHz real-time oscilloscope (Rohde&Schwarz RTP) using a second amplified photodetector. Together, the three main functional blocks (RTD, PD, VCSEL) constitute a full PRL (O/E/O) optoelectronic spiking neuron.

To enable a wide range of experimental tasks and to demonstrate the fan-in functionality (operation with more than just a single upstream optical source), two optical signal branches **A** and **B** are used as inputs to the PRL node. All of the components in these branches are coupled using fibre-optic elements. Thanks to the wide sensitivity range

of the utilised PD, the node can simultaneously process signals at various wavelengths (similarly to other O/E neurons [265]). The two input branches are identical, and each includes a tuneable laser source (at different wavelengths, namely 1310 nm and 1550 nm), an optical isolator (ISO) to limit unwanted reflections, followed by a variable optical attenuator (VOA) that allows for optical power control, and a MZM that encodes signals from an arbitrary waveform generator (AWG, Keysight M8190 12 GSa·s⁻¹) into the optical path. A 10dB RF amplification stage (Mini Circuits) is included to increase the voltage of the RF signals from the AWG (below 1V p-p) to the required levels of the MZM (between 3 to 10V). For each of the two MZMs, the applied bias voltage is set between the quadrature and the maximum of the output power transfer curve to achieve a near-linear relation between the input RF signal amplitudes and the output light intensity modulations. Finally, the two independent optical input branches are combined via a 50:50 fibre-optic coupler and fed into the input (PD) of the PRL node. Branch A is sourced with signal from a tuneable laser (Santec TSL-210) operating at the wavelength of 1310 nm, with a 10 GHz MZM amplitude modulator (Thorlabs) biased at 7.4 V. Branch B utilises a second tuneable laser (Santec WSL-110) operating at the wavelength of 1546 nm and the 10 GHz MZM amplitude modulator (JDS Uniphase) is biased at 2.1 V. The average CW optical power of the dynamical optical input signals in each branch is approximately 350 μ W.

4.4.2 Characterisation of the spiking responses in the PRL node

For a dynamical system to operate as a spiking artificial neuron, there is a set of fundamental functionalities that should be fulfilled. These include first the functionality of receiving input signals (stimuli), ideally with fan-in (allowing signals to be received from multiple upstream sources). Following that, the system needs to process these input signals, typically by the means of signal summation. These processed signals then need to be compared against a given threshold, typically with some form of leaking-over-time effect that provides the neuron with finite memory of the input signals. Finally, if the total contribution from the summed input signals crosses the spike activation threshold, the system should yield an all-or-nothing (excitable) response that is independent of

the shape of the activating perturbation(s). A single sub-threshold input perturbation (stimulus) should not result in any significant response from the system. Furthermore, an excitable spiking neuron should also act as a filter for perturbations that arrive soon after a spike is fired by inhibiting follow up spiking responses to such inputs. This is typically referred to as *refractoriness*. The following experimental results will validate all of these functionalities within the PRL node, demonstrating its ability to operate as an artificial optoelectronic spiking neuron.



As an additional remark, some of the following figures contain plots with axes labelled as "(opt.p.)". This plots depict the voltage traces (in V) as recorded at the output of the amplified photodetector (for analysis on the oscilloscope), which directly corresponds to the optical power.

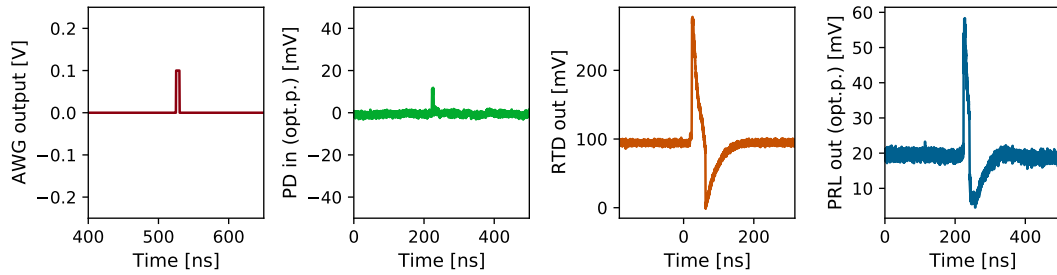


Figure 4.14: Experimental results demonstrating deterministic excitable spike firing in the O/E/O PRL node. (a) RF signal with encoded perturbation generated by the AWG which is then converted to an optical input pulse (b) using an optical modulator. (c) Electrical spike event fired by the RTD in response to the input optical perturbation. (d) Final output optical spike from the PRL node measured after the VCSEL in its structure.

First, Fig. 4.14(a) shows experimental results demonstrating the case in which a single spiking response is triggered deterministically in the PRL node. The output spiking event is triggered by a small amplitude superthreshold perturbation. This is generated in an AWG, as shown in Fig. 4.14(a), and converted to the optical domain by a MZM. This results in an optical input pulse (Fig. 4.14(b)), which is injected into the PD element of the PRL node, after which it enters the RTD. This small optical perturbation is sufficient to elicit a full, large amplitude excitable electrical response in

the RTD (Fig. 4.14(c)) that is then directly converted to an optical spike by the coupled VCSEL (Fig. 4.14(d)).

Input thresholding

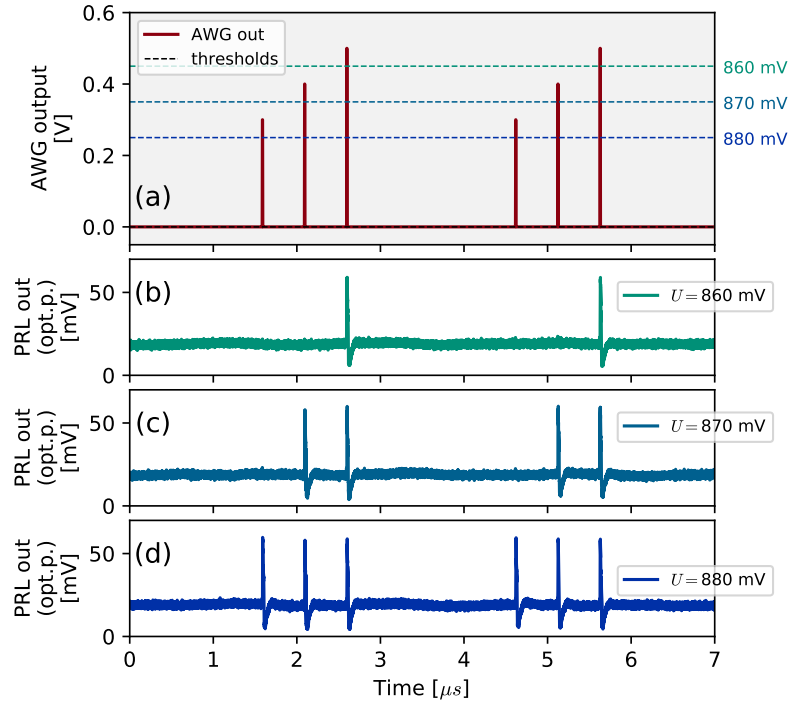


Figure 4.15: Experimental demonstration of spike firing threshold in the PRL optoelectronic spiking neuron. (a) Source RF waveform (from the AWG). Dashed lines approximate the spike firing threshold in relation to V_0 . (b)-(d) PRL node output, as recorded on an amplified photodetector. For the lowest bias voltage value (b) V_0 860 mV, only the strongest perturbation triggers a response (an optical spike). Response of the system for the same input signal but with the RTD biased now with V_0 870 mV (c) and with V_0 880 mV (d). Reproduced from [267].

First, Fig. 4.15(a-d) shows the input pulse thresholding behaviour in the optoelectronic PRL node, revealing a clearly defined threshold level. In this experiment, a ”step sequence” of three square-shaped, unipolar 3 ns pulses of gradually increasing amplitudes (mutually separated by ($t_{sep} = 500$ ns)) is encoded into one of the optical input branches of the PRL. While sub-threshold pulses do not cause any observable response at the PRL output, the super-threshold perturbation results in a fixed-shape

optical spike firing event. Furthermore, increasing the RTD DC bias V_0 value from 860 mV to 880 mV results in a shift of the spiking threshold with respect to the quiescent (steady) state of the RTD, moving the excitability threshold closer to the steady-state. In turn, this allows for spike activation for input perturbations with lower amplitudes, effectively increasing the sensitivity of the PRL optoelectronic neuron.

Refractory period demonstration

The second discussed hallmark characteristic of excitable systems is refractoriness. Once a spike has been fired by an excitable dynamical system, the latter enters a period of suppression and recovery, during which another full excitable spike can not be fired (even for super-threshold stimuli). In certain classes of artificial neuronal models, a further distinction is drawn between a so-called absolute refractory period (during which firing of responses is completely suppressed) and a relative refractory period [318] (where the probability of spike firing, and/or the amplitude of any produced spiking responses is reduced).

The experimental results demonstrating refractoriness in the optoelectronic PRL spiking node are shown in Fig. 4.16(a-c). The refractoriness test was realised using a sequence of super-threshold 3 ns-long pulses occurring in pairs (doublets) with gradually increasing mutual temporal separation. In total, $n = 15$ doublets with temporal separation values varied between $\Delta_t = 0$ ns and $\Delta_t = 150$ ns (10 ns increments) have been used. These were generated as RF signals by the AWG and encoded optically using a MZM amplitude modulation for optical injection into the PRL node. These are shown in Fig. 4.16(a). In total, twenty subsequent oscilloscope readouts (cycles) of temporal responses from the PRL node were acquired and processed into the temporal map shown in Fig. 4.16(b). Temporal maps allow us to investigate the evolution of a specific signal across repeated acquisitions. They are a useful tool for visual evaluation of arising patterns, response consistency and reliability in consecutive temporal readouts. A temporal map is synthesised by temporally aligning the studied, repeatedly occurring effects, and plotting each acquisition (cycle) as a single row of pixels in an image. The signal amplitude is represented using a colour map. There, the x-axis represents

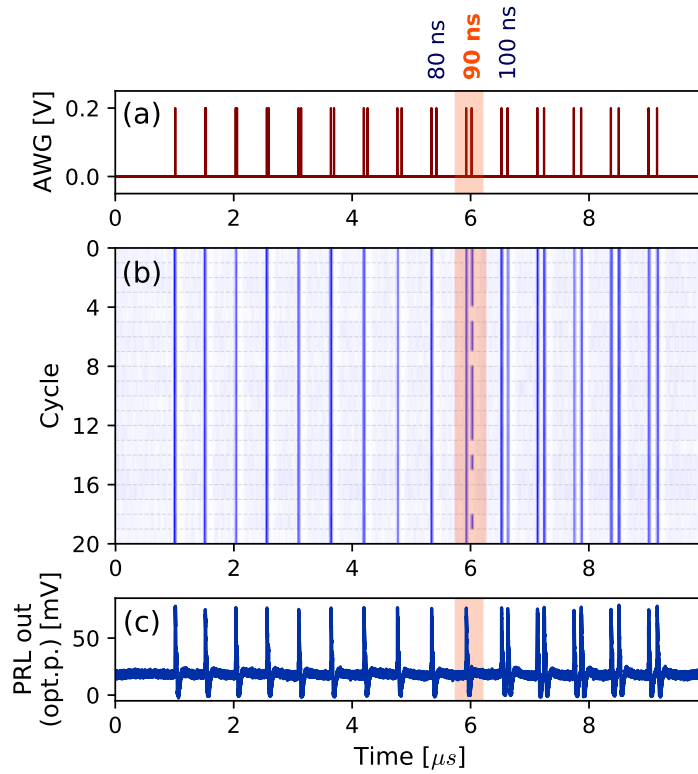


Figure 4.16: Experimental demonstration of refractoriness in the PRL spiking node: (a) Source modulation (input signal into the PRL) consisting of pulse pairs with gradually increasing temporal separation (in 10 ns increments). (b) Temporal map plotting the measured time-traces of 20 subsequently recorded measurement cycles of 10 μ s length. (c) Example of one measured time-trace from the PRL node in response to the input signal in (a). Reproduced from [267].

time and the y-axis represents a cycle number. In Fig. 4.16(b), the orange highlight corresponds to pulse doublet separation of ≈ 90 ns, for which a second excitable spiking event started occurring. Therefore, this corresponds to the absolute refractory period T_{ref} of the system. Since the upper limit of the spiking rate is governed by the refractory time, we can also estimate the maximum spiking frequency for the investigated PRL node as $\frac{1}{T_{\text{ref}}} \approx 10$ MHz. The repeated readouts in the map also show that some spikes are not fired for the interval close to the refractory period. This can be attributed to the presence of noise and temporal jitter effects in the experiment. Finally, a single cycle time trace is shown in 4.16(c). It is also worth mentioning that the temporal map in Fig. 4.16(b) confirms that the spike excitation process via super-threshold pulses

is very reliable and repeatable, with no missing first spikes observed in any of the 20 recorded cycles .

4.4.3 Pre-synaptic spike weighting in the PRL

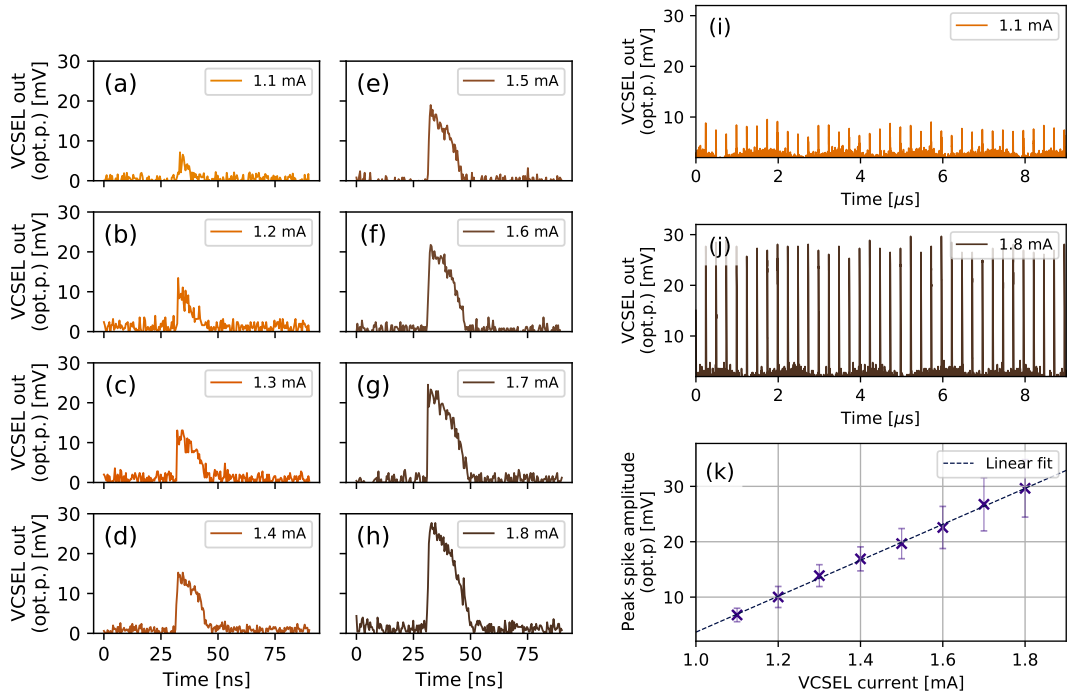


Figure 4.17: Weighting of the PRL output spike amplitude via control of the VCSEL bias current (below its lasing threshold). (a)-(h) Optical output spike amplitude when the bias current is changed from 1.1 mA to 1.8 mA with a step size of 100 μ A. In addition, individual time traces of the PRL with VCSEL biased at (i) 1.1 mA and (j) 1.8 mA. The trend is confirmed by (k) mean peak spike amplitudes over 5 repetitions as a function of the VCSEL bias, including a linear fit.

In the modular layout of the PRL, the laser element (VCSEL) is coupled to the circuit via a bias tee network. This enables for independent adjustment of the VCSEL's driving current I_{LD} , which in turn allows us to control the optical spike amplitude at the output of the PRL node. This can be considered as a case of pre-synaptic spike weighting. The lasing threshold of the VCSEL used in the experiments is approximately equal to 1.95 mA. By varying the current in a range of values below its lasing threshold (up to 1.8 mA), no light emission (lasing) is obtained from the laser (and therefore

from the PRL node) while the RTD is not spiking. However, once the RTD fires an upwards electrical spike (here triggered by the arrival of a $t_{len} = 2$ ns incoming optical perturbation), the VCSEL is momentarily brought above its lasing threshold, yielding an optical spike at the PRL node's output. Since the lasing threshold is crossed during spiking, this mode of operation can be considered as gain-switching. This brings some useful advantages: lower power consumption during the idle (non-spiking) state thanks to sub-threshold operation (no lasing), zero power level (no pedestal of constant emission) that eases the process of signal processing and weighting, and improved signal-to-noise (s/n) ratio of the produced spikes.

The experimental result demonstrating this phenomenon is shown in Fig. 4.17. Here, the VCSEL bias current was swept from 1.1 mA (Fig. 4.17(a)) to 1.8 mA (Fig. 4.17(h)) with a sweep step size of 100 μ A. As shown in the optical power traces of a single spiking event that are depicted in Fig. 4.17(a-h), the spike amplitude gradually increases during the sweep with zero (photodetector noise) floor level and well preserved shape of the spikes from the RTD. Longer time traces with series of repeated spikes are shown for the minimum (1.1 mA, Fig. 4.17(i)) and maximum (1.8 mA, Fig. 4.17(j)) investigated current. Furthermore, the trend is captured in Fig. 4.17(k). This confirms that for the used range of sub-threshold laser driving currents, a linear relation is followed for the recorded optical spike amplitude (spike peak).

This simple yet powerful function provides the PRL-class systems with an additional degree of freedom and flexibility that may prove beneficial for operation in networked arrangements, possibly decreasing the requirements of weighting elements within such networks.

4.5 Experiments with nano-pRTDs (n-pRL nodes)

4.5.1 Device, experimental setup

This section reports a novel nano-optoelectronic spiking node based on a photosensitive RTD (pRTD) element, which can also be coupled to a laser to form a nano-pRL

Table 4.5: Epi-layer stack of the nano-pRTD device. This was designed and fabricated by our project collaborators at the University of Glasgow (Dr Qusay Al-Taai, Prof Edward Wasige).

Layer	Thickness (nm)	Semiconductor	Type	Description
12	1	InGaAs	N++	Contact
11	1	InAlAs	N+	Collector
10	2.5	InGaAs	N-	Spacer
9	0.02	InGaAs	I	
8	1.7	AlAs	I	DBQW barrier
7	5.7	InGaAs	I	Well
6	1.7	AlAs	I	DBQW barrier
5	0.02	InGaAs	I	
4	0.2	InGaAs	N-	Spacer
3	1	InAlAs	N+	Emitter
2	5	InGaAs	N++	Contact
1	1	InAlAs	I	

architecture. In this layout, the optical input functionality (previously provided by a commercially-sourced amplified photodetector coupled to an RTD) is now directly provided by a photosensitive window in the fabricated RTD structure. This permits us to deterministically generate spike-firing events in the new RTD design using direct injection of light signals. The pRTD therefore represents a step forward towards miniaturised optoelectronic spiking RTD nodes. Furthermore, these pRTD devices employ a new type of nanoscale upper electrode (previously also referred to as a nanoinjector [277]) to decrease the currents flowing through the device. Therefore, these devices will be referred to as nano-pRTDs, or n-pRTDs. A manuscript describing this device and its dynamical electronic behaviour was recently published in [319].

These new InGaAs/AlAs DBQW nano-pRTD devices have been fabricated by our ChipAI EU project collaborators at the University of Glasgow (Prof Edward Wasige's group). The epilayer stack is shown in Table 4.5. A simplified scheme of the device, showing the upper nanoinjector electrode, and SEM images of the device and the nanopillar are shown in Fig. 4.18. The final die (chip) containing many fabricated pRTD devices with bond pads is shown in Fig. 4.19.

Thanks to the optical sensing capability of the pRTD, the coupled external photodetector is no longer required. This significantly simplifies the experimental layout. A

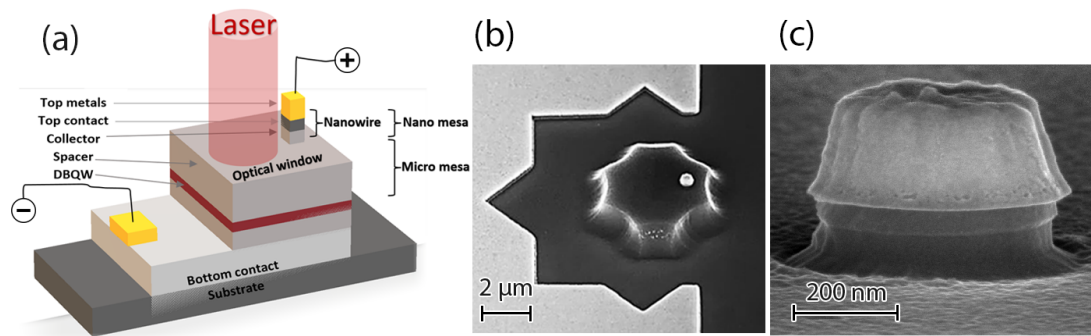


Figure 4.18: (a) Scheme of the photodetecting RTD with nano-injector (nano-pRTD). (b,c) SEM images of the upper, nanoscale electrode of the nano-pRTD. Reproduced from [319].

simple scheme describing the new experimental setup is shown in Fig. 4.20. For the new n-pRL node, the optical input signals are directly launched into the embedded optical window of the nanostructure RTD using a lensed-end optical fibre. The GSG RF probe provides the DC bias that sets the operation point V_0 of the RTD, and collects the output (spiking) signals from the pRTD via the RF port of the bias tee. All the used circuit component models remain the same as introduced in the previous (PRL) section.

The I - V characteristic of the nano-pRTD device is shown in Fig. 4.21. The device exhibits a pronounced NDR region (highlighted in orange) from approx. 590 mV (reverse biased) and extending towards 735 mV. The peak-to-valley current ratio (PVCR, dark) is approx ≈ 1.38 . Thanks to the optical sensing property of this device, we can directly investigate the effect of external illumination with infrared (IR) light on the shape of the I - V curve. The I - V curve in Fig. 4.21 is recorded both without illumination (dark) and with 1 mW of CW IR laser light illumination ($\lambda = 1310$ nm) directed on the RTD structure using a lensed-end optical fibre (OZ Optics). Under external IR light illumination, Fig. 4.21 reveals a change in the shape and features of the I - V curve. First, there is an additionally generated photocurrent (with estimated spectral responsivity of ≈ 1.09 A/W at the given wavelength), as well as a small (≈ 6 mV) shift of the peak voltage towards lower values.

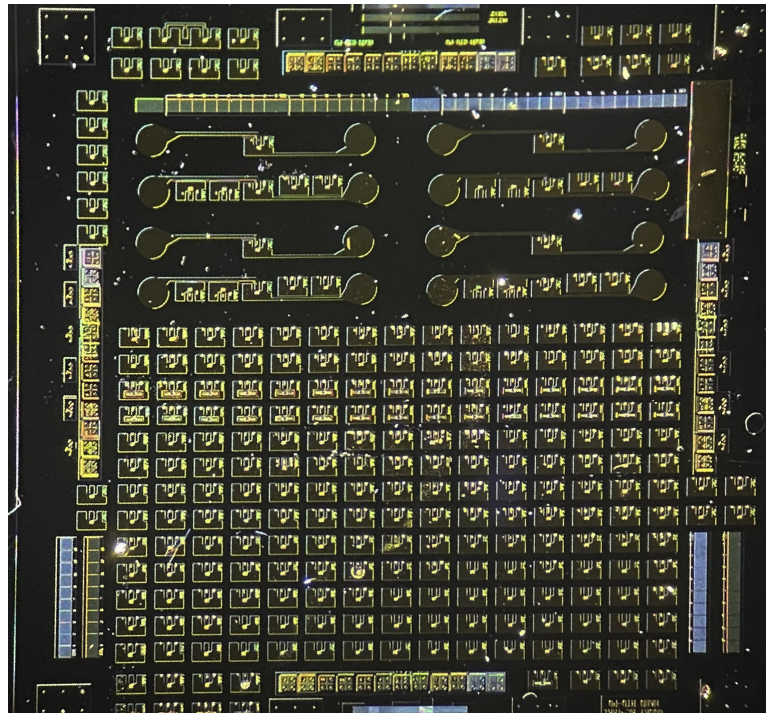


Figure 4.19: Photograph of the die containing the pRTD devices. This sample was fabricated Dr Qusay R. A. Al-Taai and team of Prof. Edward Wasige at the University of Glasgow.

4.5.2 Dynamical characterisation: pRTD

The dynamical characterisation for the nano-pRTD device follows the same steps as previously discussed in Sec. 4.4.2. A single nano-pRTD device was driven using a RF probe (Picoprobe) connected to a bias tee (Inmet). The latter provides the DC bias V_0 to the nano-pRTD device, with the RF terminal providing connectivity to the laser diode (here, a VCSEL) that converts the elicited spikes in the system to the optical domain. The nano-pRTD is subject to external injection of optical input signals directly coupled to the light-absorptive window of the device. The optical input coupling was maximized by monitoring the photocurrent contribution and maximizing the current flowing through the device for fixed values of pRTD bias voltage, V_0 and optical power P_0 . This was done by finely adjusting the positioning and focusing of the lensed-end fibre used for optical coupling. The dynamical optical input signals used for analysis were provided in the same way as discussed in Sec. 4.4, from a 1310 nm tuneable laser

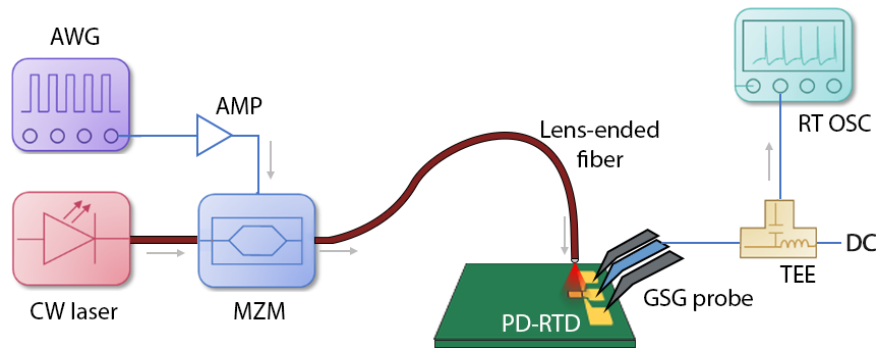


Figure 4.20: Simplified scheme of the experimental setup used to operate the nano-pRTD spiking neuron.

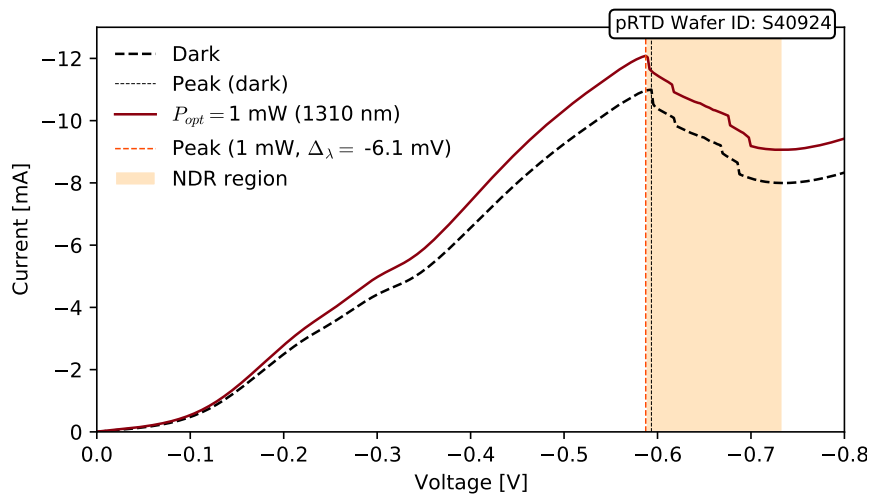


Figure 4.21: I - V characteristic of a solitary nano-pRTD device (reverse biased) under darkness (black dashed line) and when subject to 1 mW CW IR laser light illumination (solid red line).

(Santek).

First, the dynamical response of the system was evaluated independently using only the nano-pRTD, without connecting the VCSEL as an optical output for E/O conversion. The experimental dynamical analysis results obtained for a nano-pRTD device operating in reverse bias conditions are shown in Fig. 4.22. This figure provides an overview of all the fundamental spiking characteristics for both the peak ($V_0 = 590$ mV) and valley ($V_0 = 688$ mV) biasing conditions. The recorded time traces in Fig. 4.22 clearly confirm that three key dynamical characteristics are observed for both

(peak and valley) operational points:

1. deterministic and reliable spike triggering via super-threshold perturbations (Fig. 4.22(a,b)),
2. all-or-nothing character of the elicited spiking responses with well-defined excitability threshold ((Fig. 4.22(c,d)),
3. existence of absolute refractory period (Fig. 4.22(e,f)), during which the system cannot fire a new (second) spike after a first excitation.

Notably, the results in Fig. 4.22 also demonstrate how the character of the nano-pRTD response differs when the device is operated in the peak and valley points. These two cases can, to a certain extent, be considered as opposites (inverses) of each other. Spikes fired at the peak have a pronounced main part (here aiming downwards due to reverse biased operation of the nano-pRTD) with a small "bump" in the opposite direction (upwards). On the other hand, the spikes elicited when operated in the valley follow an opposite trajectory, following an upwards-downwards direction. In this case of valley operation, the upwards and downwards excursions of the fired spikes are almost symmetrical and exhibit a large peak-to-peak amplitude (voltage) ratio. Furthermore, the spiking at the I - V valley operation point exhibits a shorter refractory period ($t_{\text{refV}} \approx 70$ ns) compared to the case of biasing and operation at the peak ($t_{\text{refP}} \approx 120$ ns), rendering the valley spiking dynamics almost 40% faster. In addition, spikes elicited at the valley exhibit much more pronounced relaxation oscillations once the spike is fired. This is likely caused by the shallow shape of the valley I - V characteristic compared to the very sharp transition between the first PDR and the NDR at the peak (Fig. 4.21). This dynamical behaviour has some interesting consequences for processing sequential signals, as can be seen in the sudden spiking suppression in the orange-highlighted area in Fig. 4.22(f). This phenomenon will be further studied and reported in Sec. 4.5.4.

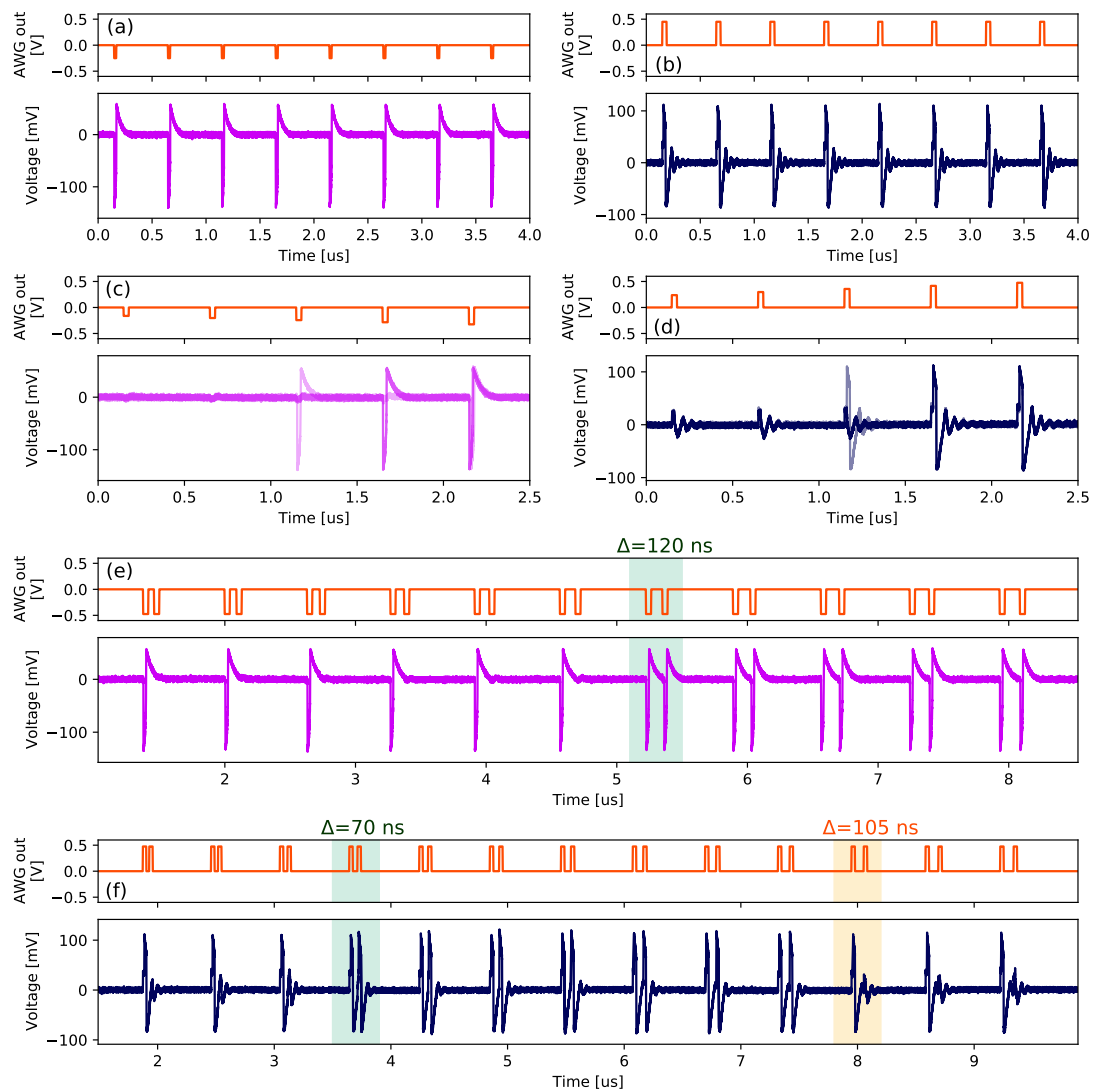


Figure 4.22: Characterisation of the spiking dynamical behaviour of the nano-pRTD node when biased at the peak ($V_0 = 590$ mV) and valley ($V_0 = 688$ mV) operation points. Repeated deterministic spike triggering elicited in the nano-pRTD when biased at the peak (a) and the valley points (b). Spike thresholding behaviour of the nano-pRTD when biased at peak (c) and valley points (d). Refractory period of the nano-pRTD when biased at the peak (e) and valley points (f). We must note that in order to match the corresponding operational modes at the peak and valley biasing points, the optical input trigger signals (generated from the AWG) are designed with negative/positive polarity.

4.5.3 Dynamical characterisation: coupled spiking nano-pRTD-LD

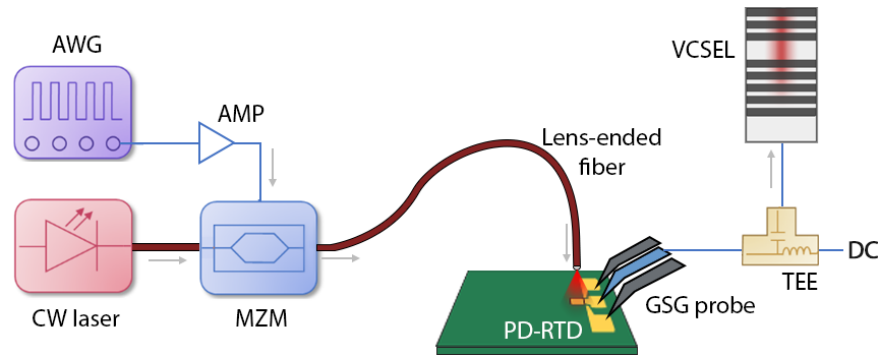


Figure 4.23: Simplified scheme of the experimental setup used to operate the nano-pRL optoelectronic spiking neuron. In comparison to Fig. 4.20, there is a laser connected in series at the output.

The dynamical characterisation of independent pRTD elements is now followed by the analysis of systems with both input and output (I/O) enabled for optical signals. In a similar fashion to the previous PRL case powered by a micrometre-scale RTD (Sec. 4.4), the nano-pRTD can also be coupled to an off-chip light-source (such as a laser diode, LD) to enable E/O signal conversion at the output. Such a system can be referred to as nanoscale pRTD + LD, or nano-pRL in short. Here, the architecture is realised by coupling a nano-pRTD to a telecom-wavelength VCSEL. In this case, as shown in Fig. 4.23, the RF port of the bias tee connects directly to a VCSEL (via the RF port of the laser mount) for conversion of nano-pRTD spikes into the optical domain. Such a coupled layout represents an O/E/O-enabled system with both input and output optical terminals and is therefore of interest for the development of spiking optoelectronic nodes in future neuromorphic photonic platforms.

Here, the same spiking characterisation steps are followed as in Sec. 4.4. This is a nano-pRL device from the same wafer as in Sec. 4.5.2, and is again operated in reverse bias. The experimental results on the characterisation of the spiking dynamical responses in the nano-pRL nodes are shown in Fig. 4.24 (thresholding) and Fig. 4.25 (refractoriness). In all these plots, the smaller green-coloured time-traces at the top

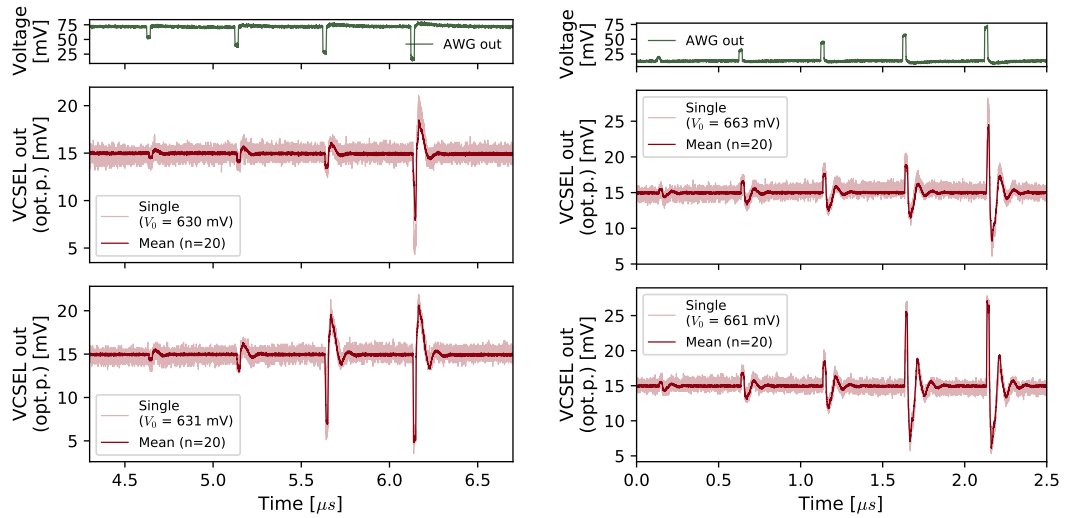


Figure 4.24: Evaluation of thresholding in the nano-pRL system. Green time traces show the input RF modulation waveforms, which are used (via the MZM) to modulate CW IR light signal input to the nano-pRL system, while read time traces show the optical output from the VCSEL in the nano-pRL node. Left column shows responses in peak operation points, while right column shows responses in the valley operation points. In both cases, two different RTD bias points (altering the distance of the steady state from the spiking threshold) have been evaluated.

represent the modulation RF signals used as amplitude modulation for CW IR light injected into the photo-absorptive window of the nano-pRTD. The light red traces show individual repeated nano-pRL readouts overlaid together ($n = 20$), and the dark red trace represents a mean trace of these repeated measurements. The nano-pRL node was probed both at the peak ($V_0 = 630$ mV) and at the valley operation point ($V_0 = 663$ mV). First, Fig. 4.24 demonstrates the optical spiking and spike thresholding capability of the nano-pRL system for both peak (left column) and valley (right column) operation points. Both of these operational points yielded deterministic optical spiking regimes at the nano-pRL output (coupled VCSEL) in response to optical input perturbations. Furthermore, Fig. 4.24 shows how controlling the applied RTD DC bias V_0 allows us to tune the distance from the spike firing threshold in the system. By increasing V_0 by 1 mV to 631 mV, the relative distance between the steady-state and spiking decreases, and spikes are elicited for a lower amplitude optical input perturbation (second to last) as well. Similar threshold shifting steps were performed when the system was

operated in the valley operation regime ($V_0 = 663 \text{ mV} \rightarrow 661 \text{ mV}$). In the valley, the NDR is approached from the 2nd PDR region (higher voltage values), therefore the threshold is moved closer to the steady state by reducing the applied RTD bias V_0 . Again, shifting the threshold closer activates an additional spike, now triggered by a previously sub-threshold perturbation.

In the valley operation point, there is a more pronounced input perturbation amplification prior to the activation of spiking responses. This effect has been observed in certain devices, and can be attributed to the amplifying nature of the RTD within the NDR region. The valley-triggered spikes also exhibit more pronounced relaxation oscillations after the spike is fired. This can be attributed to the shape of the I - V and is likely described by the "shallow" character of the I - V in the valley operation point (comparably lower values of both first and second derivative of $f(V)$) around the inflexion point between the regions of NDR and 2nd PDR).

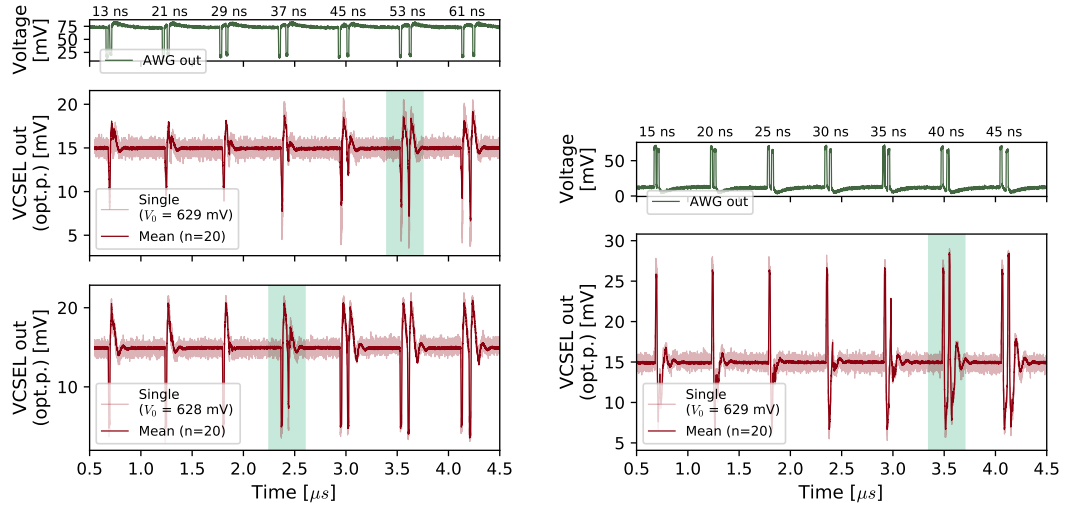


Figure 4.25: Evaluation of refractoriness in the nano-pRL system. Green time traces show the input RF modulation waveforms, which are used (via the MZM) to modulate CW IR light signal input to the nano-pRL system, while read time traces show the optical output from the VCSEL in the nano-pRL node. Left column shows responses in the peak operation point ($V_0 = 630 \text{ mV}$), while right column shows responses in the valley operation point ($V_0 = 663 \text{ mV}$).

Second, the refractory times for both operational points have also been experimentally evaluated. The results for this experimental analysis are shown in Fig. 4.25. The absolute

refractory period estimation for peak operation point is around $t_{\text{refP}} \approx 50$ ns, while for the valley biasing point, it is estimated to be approx. $t_{\text{refV}} \approx 40$ ns. Notably, Fig. 4.25 also demonstrates that, for the case of valley biasing, a relative refractory period can also be observed, with a second spike of lower amplitude occurring at a trigger pulse spacing of 35 ns. For higher trigger separation between the two stimuli, two full-fledged consecutive spikes are fired, as expected for perturbations with larger separation than t_{refV} . In agreement with previous measurements, we observe that the spiking regimes achieved in the valley are faster (in this case, about 20%), and that these exhibit more pronounced relaxation oscillations. The overall increased dynamics speed (as compared to the case of the solitary pRTD) are likely to be caused by variations in the experimental electronic setup between these two cases, including the signal coupling to the RF-coupled laser mount used to drive the VCSEL.

4.5.4 Resonate-and-fire behaviour

Previous sections have highlighted the occurrence of pronounced decaying oscillations when the (valley biased) RTD circuit relaxes after receiving incoming perturbations, including also after the firing of an excitable spike. From a dynamical standpoint, these oscillations can be viewed as dampened harmonic oscillations attracted towards the fixed (valley) point in the I - V curve. These oscillations are more pronounced when the system is operated in the valley, thanks to the wider, more shallow basin of attraction near the biasing point. Interestingly, thanks to these relaxation oscillatory dynamics, a state of the dynamical system is, among other factors, a function of what happened in the system in a short interval prior. This can be considered as a form of a short-term memory, which may hold potential for processing of (temporal) information in the RTD. Dampened oscillations during system relaxation have a direct effect on the way excitable responses are elicited in dynamical systems. Unlike in the more traditional leaky integrate-and-fire (LIF) neuronal models (Sec. 1.2.4), where the internal state variable decays exponentially after every (sub-threshold) perturbation (Fig. 1.3), in these types of neurons (called resonate-and-fire, R&F neurons), the state variable decays via dampened harmonics.

This resonant behaviour fundamentally changes the responses of the RTD neuron to temporally-spaced perturbations. One such novel behaviour comes in the form of additional inhibitory effects. Classical LIF neurons are functionally comparable to low-pass filters (on input perturbation repetition rates). Any two sufficiently strong pulses temporally separated by a time interval of at least the refractory period will trigger two spiking responses. On the contrary, R&F neurons behave more like band-pass frequency filters - responses to input stimuli (those that are temporally separated above the refractory period) can be inhibited due to the oscillatory effects. If a second input stimulus arrives in the lowest point (antinode) of the relaxation, a second spike is not elicited.

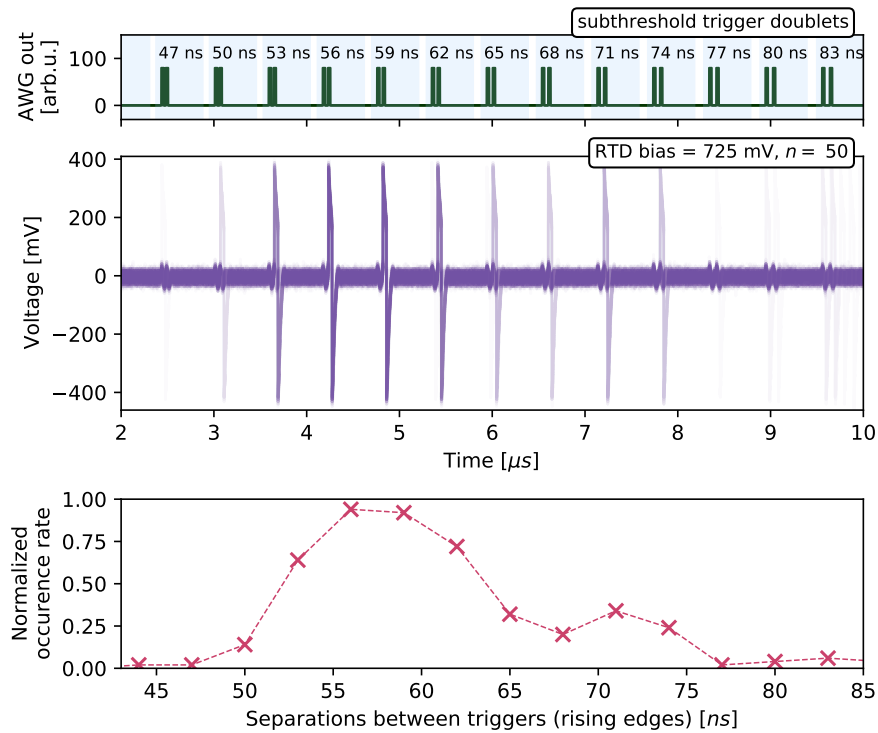


Figure 4.26: Resonate-and-fire effect in the pRTD. The green time trace (top) shows the modulation waveform (used in MZM for optical input into the pRTD) with pairs of sub-threshold perturbations. Purple (middle) trace shows overlay of $n = 50$ pRTD electrical output time traces. Dashed (pink) bottom time trace shows the normalised occurrence rate of pRTD spiking for each pair of input sub-threshold perturbations.

The dampened harmonic oscillations can also help elicit spiking from sub-threshold stimuli at specific "resonant" rates. This phenomenon is demonstrated in Fig. 4.26, which

shows an overlay of $n = 50$ repeated acquisitions. In this experiment, a valley operated, reverse-biased pRTD device was subject to directly coupled optical perturbations in the form of pulse doublets of gradually increasing temporal spacing. This is a test similar to the refractory period evaluation. However, unlike in the refractoriness evaluation, both of the optical input stimuli in this experiment have an amplitude (closely) below threshold. For a system without any form of local memory, no spikes should ever be fired. However, for a resonate-and-fire system, a spike should be fired for doublets with spacing equivalent to the resonant repetition rate. Both lower and larger spacing should yield no excitable spikes.

Fig. 4.26 shows the occurrence probability plot, confirming the resonant firing behaviour in the pRTD device. Optical perturbation doublets with mutual temporal separation lower than 50 ns trigger no spikes in the pRTD. However, as the temporal separation further increases, spikes are suddenly being elicited despite the fact that the trigger perturbations are sub-threshold stimuli. This effect peaks at a "resonant rate" of $T_{\text{reso}} \approx 57$ ns, where the measured spike firing probability reaches a value very close to unity (meaning that the resonant spike firing effect is very reliable). As the temporal spacing further increases, the spike firing probability drops again, with a second smaller increase of the spike firing probability at $T_{\text{reso}} \approx 71$ ns. For longer separations, the circuit completely stabilises and the dampened oscillations vanish. As a result, no resonant spike firing effect occurs and the spike occurrence probabilities approach zero. These results confirm, that the (valley-operated) pRTD device operates as resonate-and-fire spiking neuron with strongly dampened oscillations.

Chapter 5

Applications and networks of RTD-powered neuromorphic nodes

Relevant authored peer-reviewed manuscripts:

- ◆ **Hejda, M.**, Alanis, J.A., Ortega-Piwonka, I., Lourenço, J., Figueiredo, J., Javaloyes, J., Romeira, B. & Hurtado, A. *Resonant Tunneling Diode Nano-Optoelectronic Excitable Nodes for Neuromorphic Spike-Based Information Processing*. Phys. Rev. Applied 17, 024072. [10.1103/PhysRevApplied.17.024072](https://doi.org/10.1103/PhysRevApplied.17.024072) (2022)
- ◆ **Hejda, M.**, Malysheva, E., Owen-Newns, D., Al-Taai, Q.R.A., Zhang, W., Ortega-Piwonka, I., Javaloyes, J., Wasige, E., Dolores-Calzadilla, V., Figueiredo, J.M.L., Romeira, B. & Hurtado, A. *Artificial optoelectronic spiking neuron based on a resonant tunnelling diode coupled to a vertical cavity surface emitting laser*. Nanophotonics 0362. [10.1515/nanoph-2022-0362](https://doi.org/10.1515/nanoph-2022-0362) (2022)

WITHIN this chapter, the prospects of resonant tunnelling diode (RTD) devices for information processing will be investigated. Since generalised neuromorphic systems represent a specific type of a computer (or computing architecture), a demonstration of data and information processing based on the neuron-like spiking regimes exhibited by the RTDs is of significant interest. This chapter aims to go beyond focusing on exploration of the phenomena of excitability in these devices, and extends the scope of using the high-speed spiking dynamics of RTDs for various methods of information

encoding, handling, and processing. These methods are considered from two different standpoints:

- processing of signals and data by individual RTD devices, utilising their nonlinear (spiking) dynamics for computation (Sec. 5.1, Sec. 5.3), and
- processing of signals and data by networked arrangements of multiple RTDs acting as spiking nodes (Sec. 5.2, Sec. 5.4).

Both theoretical and experimental studies are reported. The numerical results demonstrated in this Chapter are realised using the dynamical O/E/O RTD models introduced in Sec. 4.3. Experimental demonstration of information processing tasks is provided for the case of a single PRL node in Sec. 5.3, reporting dual input coincidence detection and exclusive OR (XOR) task respectively, and for the case of a multi-RTD functional interlinking in Sec. 5.4.

5.1 Information processing in a numerical PRL-MN node

5.1.1 Time-domain multiplexed pattern recognition

As was discussed in Chapter 4, the dynamical state of an RTD at any given point in time not only depends on the current conditions, but also on perturbations entering the system in short period of time prior. Therefore, if data is provided to the node in a temporally sequential manner (such as using a pulse amplitude modulated (PAM) sequence, a spike train or an analog RF signal), it can be said that the nonlinear dynamical system will process these inputs while exhibiting a rolling, short-term memory.

Here, this capability of a photodiode-RTD-laser monolithic node (PRL-MN, Sec. 4.3) is numerically demonstrated and utilised to recognise small (2×3 pixel) binary patterns characterising different pieces in the game Tetris. The results of this procedure are shown in Fig. 5.1. In total, four Tetris piece patterns are investigated: a 2×2 square; 1×3 vertical line; L-shaped piece and a diagonal piece. To achieve recognition of an arbitrary selected pattern, a pre-processing step is applied. Without the pre-processing,

the PRL-MN would only fire in the singular case of six sequential True values. Therefore, any pattern that can be uniquely remapped to such sequence will be recognised by the system. In this experiment, the diagonal Tetris block (fourth in the order, shown on the right hand side of the figure) has been arbitrarily selected as the target pattern, and a custom binary multiplier matrix (kernel) has been designed to perform the remapping for recognition of this pattern. The binary kernel is shown at the top of Fig. 5.1 in green/orange and remains fixed all through the performed numerical experiment. The pre-processing code adjusts each pixel value from the source pattern using the following code:

```

1 def Kernel_Op(d,k):
2     if d == 1:
3         if k == 1:
4             return 1
5         elif k == 0:
6             return 0
7     elif d == 0:
8         if k == 1:
9             return 0
10        elif k == 0:
11            return 1

```

Because a single RTD-MN can only process data in the form of a time series, the pre-processed 2D Tetris pattern had to be flattened* from 2D to 1D vector of binary values. The pre-processed, serialised vector was then turned into an electronic modulation for the RTD-MN. Each value of the binary vector was encoded into the amplitude of a single, RZ-encoded sub-threshold square pulse with amplitude of $V_{\text{pulse}} = -17 \text{ mV}$, 50 ps length and 1:1 ratio (50% duty cycle) for the PAM RZ scheme. In total, up to six pulses represent the input Tetris-like pattern. Different pre-processed pulse sequences and their corresponding input Tetris pieces are shown in Fig. 5.1(a). In this demonstration, no light input is provided to the RTD-MN, which is denoted in dashed purple line for input optical modulation. The RTD in this RTD-MN node was biased with $V_0 = 730 \text{ mV}$ (with a corresponding current of $I_{DC} = 73 \text{ }\mu\text{A}$), thus operating in the valley region of

*using standard numpy function `array.reshape()`

the I - V . The nano-scale laser in the model was provided with an additional biasing current term $J = 280 \mu\text{A}$ (in close proximity but below its lasing threshold current).

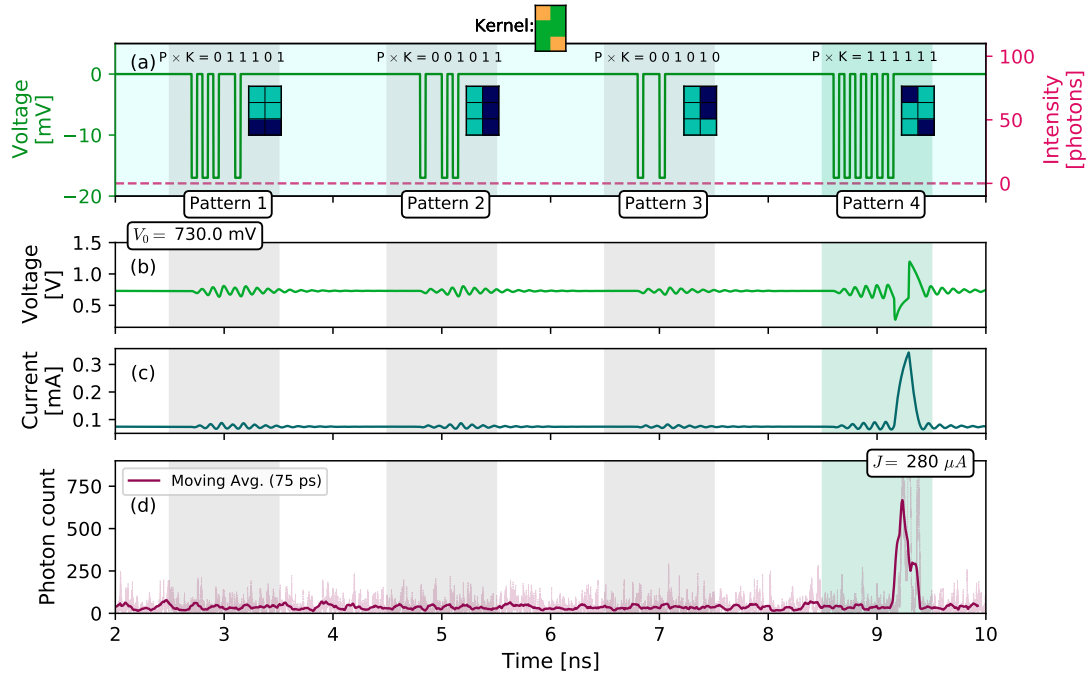


Figure 5.1: Sequential data processing for the recognition of serialised Tetris-like input patterns, realised using the PRL-MN spiking optoelectronic neuron model. Only for the case where the input pattern P matches the kernel K , the pattern is pre-processed into a sequence of input pulses that triggers an excitable spike in the RTD. (a) Input pulse sequences. (b) PRL-MN voltage response. (c) PRL-MN current response. (d) PRL-MN optical response.

The results for the system responses to each of the four pre-processed, serialised Tetris patterns are shown in Fig. 5.1(b-d). For any of the non-target patterns (square, vertical, L-shape Tetris blocks), the system does not produce a spike, and only the desired (diagonal) Tetris pattern remaps to a complete sequence of six pulses whose combined effect adds up (integrates) and elicits a full, excitable response in the PRL-MN. This can be observed both in the electrical and optical (Fig. 5.1(d)) output traces. By accounting for the refractory period of this system, the estimate for achievable processing speed for such 6-bit binary patterns in a single node is around 1.1 ns/pattern ($6 \times 2 \times 50 \text{ ns} + 500 \text{ ps}$ of refractory period estimate). The excitable response for the desired six pulse sequence also exhibits very good repeatability, with spike being elicited very

reliably in the model. Therefore, this demonstrates that even a single optoelectronic RTD-MN spiking neuronal node can be utilised for the recognition of high-speed binary bit sequences.

5.1.2 Sequential convolution-like operation for feature detection

Furthermore, the sequential bit sequence processing functionality in combination with the pre-processing (pre-weighting) can be utilised to perform certain image processing tasks from the field of computer vision. In one such task, the goal is to detect patterns occurring in images. This is typically realised by performing a discrete convolution, which represents one of the most important algorithms in computer vision as it allows for spatially-aware processing in multidimensional image data. The core functionality of such algorithm relies on multiplying a sub-matrix from the source (image) with a smaller matrix (typically called a convolution kernel), followed by accumulating the product. This procedure is therefore referred to as multiply-and-accumulate, or MAC operation. The larger source image is scanned through (either step-by-step or with larger step defined as a *stride*) to obtain the sub-matrices using a principle of sliding (the sub-matrix is sometimes referred to as a sliding window).

Here, the single RTD-MN node is demonstrated as being suitable for realisation of the accumulate functionality for the MAC operation that underlies the process of the discrete convolution for image feature extraction. Similarly to the previously shown Tetris piece recognition task, this is enabled by the temporally integration functionality observed in the RTD spiking node.

The main workflow of this method is shown in Fig. 5.2(a) and directly reproduced from one of our previously published manuscripts [320]. A square binary image M of size $n \times n$, in this case binarized as $(-1,1)$, serves as an input for this task. The image is processed on a row-by-row basis, where during the pre-processing step, an element-wise product between a $n_k \times n_k$ matrix kernel K with $n_k = 3$ and sub-matrix of the binary image M_h is pre-computed as:

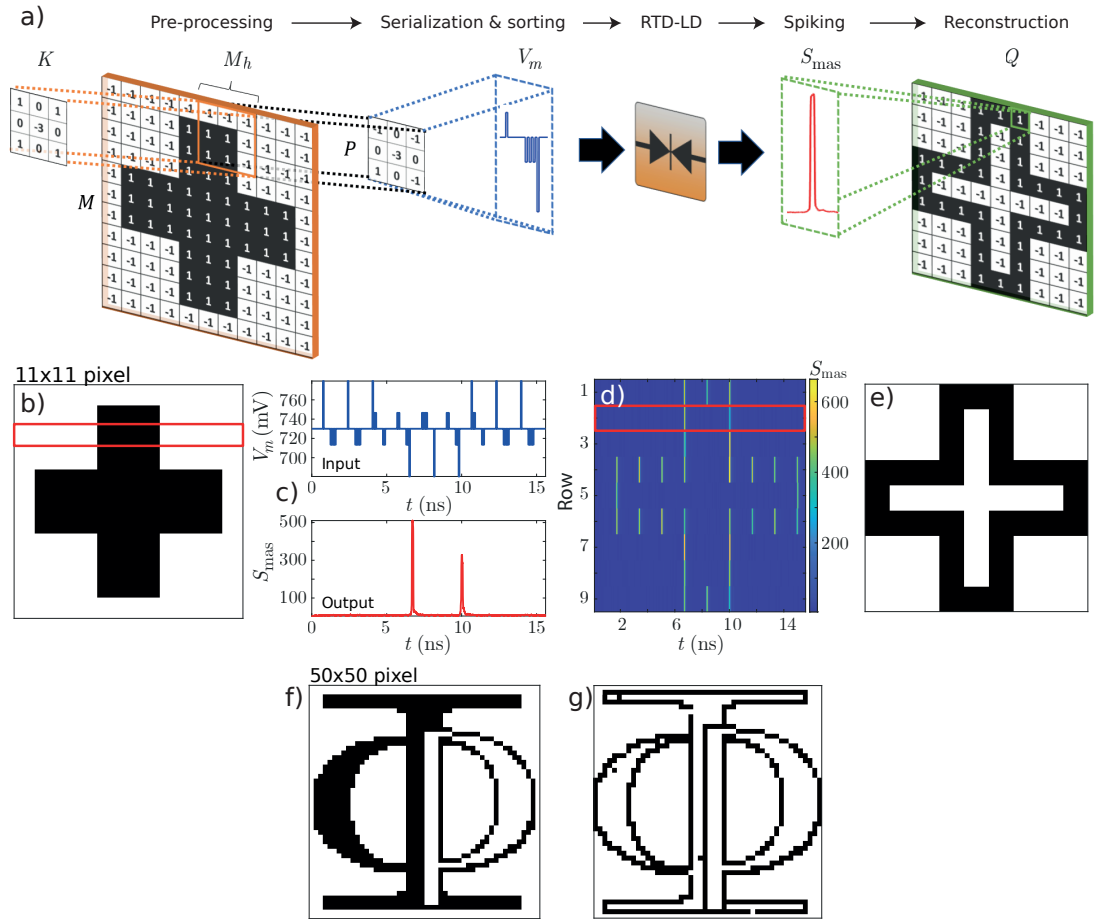


Figure 5.2: Processing steps and example of the convolution-like feature extraction (edge detection) in the PRL-MN. a) Schematic showing the individual steps of data processing in the PRL-MN model. A sliding window approach is utilised, followed by a MAC operation for each window position, serialisation, sorting and dual-polarity PAM modulation. This is fed into the optoelectronic RTD model, which may respond with optical spikes. b)-d) Shows individual steps for processing (edge feature detection) for an example of a black-and-white cross image. e) Cross image with extracted features. f) Second example, using larger image with Institute of Photonics (IOP) logo. g) IOP logo with features extracted using the same process. Reproduced from [320].

$$P = K \circ M_h = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -3 & 0 \\ 1 & 0 & 1 \end{bmatrix} \circ \begin{bmatrix} M_{i,j} & M_{i,j+1} & M_{i,j+2} \\ M_{i+1,j} & M_{i+1,j+1} & M_{i+1,j+2} \\ M_{i+2,j} & M_{i+2,j+1} & M_{i+2,j+2} \end{bmatrix} \quad (5.1)$$

where i and j are the individual pixel indices within M_h . The resulting pre-processed

3×3 matrix P is reshaped[†] into a one-dimensional (9×1) vector and sorted in descending order, ensuring that all negative pulses are integrated consecutively to elicit a firing response. This vector is then time-domain multiplexed into a modulation RF signal for injection into the RTD-MN node. Each vector element is encoded into the amplitude (voltage) of a single, return-to-zero (RZ), 50 ps long square pulse with 50 ps separation (corresponding to 1:1 RZ with total of 100 ps per bit). Therefore, this can be considered as a case of a PAM scheme. The amplitude is assigned as $V_{\text{pulse}} = 16 \text{ mV} \cdot P_i$. Once the current image sub-matrix is processed, the sliding window moves within the row with stride = 1. All the processing steps within a single row are concatenated (time domain multiplexed) into a single RF time trace, which is provided as an electrical input into the PRL-MN. The process described above is repeated for $n - n_k + 1$ rows to account for all rows of the source image M . Finally, the output binary image Q is reconstructed based on the output time traces, using spiking events and quiescent responses as the two binary states for the black and white pixels.

The procedure described above is utilised to process an 11×11 binary image with a 'cross' pattern (Fig. 5.2) and used to demonstrate feature extraction (edge detection). Values from each row triplet (a selected top row of the triplet shown in red in Fig. 5.2) in M are pre-processed and RZ (1:1) encoded into V_m as shown in the example in the upper part of Fig. 5.2(c). In total, there are 9 time-multiplexed patterns ($n = 11$, kernel $n_k = 3$ reduces by $n_k - 1$) with a duration of 9×100 ps per pattern and temporally separated by 750 ps to account for the time refractory period. The RTD within the node was biased in the valley region with a voltage of $V_0 = 730$ mV. The corresponding optical output of the PRL-MN node is shown as S_{mas} at the bottom of Fig. 5.2(c). The optical output shows clear occurrence of two spikes, triggered by two distinct patterns (pixels 4 and 6). Output traces for all the processed rows are shown in Fig. 5.2(d), with colour coding for the local PRL-MN optical output. This figure clearly reveals that occurrences of optical responses in S_{mas} correspond to a detected feature (edge) in the source image. This is further visually shown by converting the spike-map into a binarized image representation of the PRL-MN de-multiplexed output data Fig. 5.2(e),

[†]using standard numpy function `array.reshape()`

which confirms that this processing method has perfectly captured all existing edge information from the source image.

To further demonstrate and validate the capability of this RTD-enabled image feature extraction method, the same processing workflow is utilised on a larger 50×50 binary image representing the University of Strathclyde's Institute of Photonics crest (Fig. 5.2(f)). The processed and temporally demultiplexed PRL-MN output data (visualised as a binarized image) show that the excitable spiking functionality can be paired with a simple pre-processing algorithm to achieve highly reliable (99.7% accuracy in case of the 50×50 image) feature extraction directly in the spiking domain and at very high ($> \text{GHz}$) speeds.

5.2 Spiking optoelectronic neural network with RTD-MN nodes

The information processing capability of ANNs and ML systems can be assumed to directly correlate (increase) with the overall system scale and complexity. For this reason, scaling up remains the most crucial (and also arguably the most challenging) aspect for neuromorphic photonics systems. While certain architectures and implementations may provide sufficient computational advantage with low node count implementations (such as time-delayed reservoirs), it is likely that practically useful neuromorphic chips will require an increased number of artificial neurons (realised by physical devices) operating independently and forming a physical neural network. Therefore, demonstration of networked arrangements is of key importance for prospective neuromorphic candidate devices such as the RTDs.

To validate a networked mode of operation for the RTD-powered spiking nodes, a photonic spiking neural network (pSNN) is numerically realised utilising the physical RTD-MN models (dynamical RTD equations) as network nodes. The studied networked arrangement is n -to-1, which directly allows for n -to- n parallelisation without additional model extension. These arrangements can be considered as a spike-operating variation on the single layer, feed-forward perceptron model with all-to-one layout. Therefore,

within the remainder of this work, this model will be referred to as RTD-SpikeSLP (spiking single layer perceptron). Such a network has a single layer of excitable nodes that process input analog signals or pulse-encoded data. The output from each node in this layer is coupled to a weightable, uni-directional (feed-forward) optical link that propagates the signal towards the single downstream node and performs the weighting operation (such as amplitude or power adjustment) on the excitable spikes. Finally, all of the links undergo a fan-in and all the signals come together. These are provided as an input to the photo-detecting part of a single downstream PRL-MN node which performs signal summation. In case the weighted input sum of incoming signals exceeds the firing threshold (given by the PRL-MN bias V_0 and photodetector conversion efficiency among other factors), another spike is fired by the output (downstream) node. This represents the output of the spiking network. We have published this approach in a recent publication [320]. In this demonstration, the network is utilised in a feed-forward

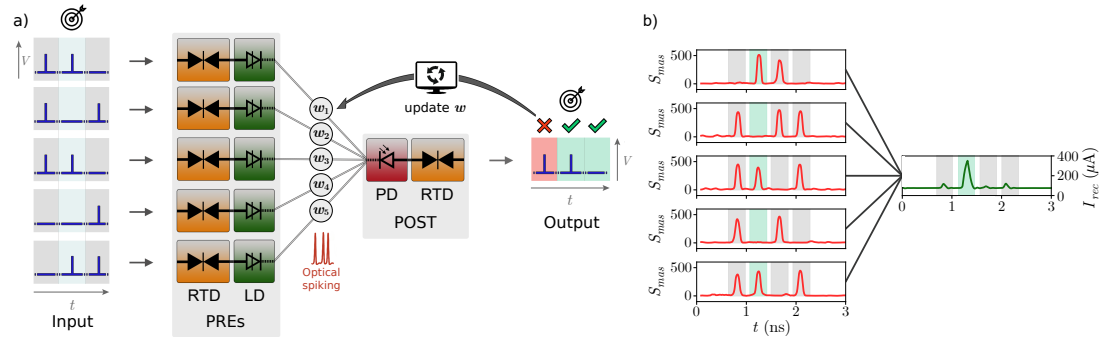


Figure 5.3: *a)* Network architecture diagram. The output state of the downstream node is compared to the label, and if there is a mismatch between label and output state, the weights are updated. The desired pattern is highlighted with the target icon. *b)* visualisation of inference in the 5-to-1 feed-forward network model. The guiding signals representing pattern labels are visualised as background shading (green for 'True', grey for 'False'). Only a particular spatial pattern ([1 0 1 0 1], green) results in firing of an electric spike by the downstream RTD-PD node (green current trace). The red time trace represents a simple moving average of the LD output optical signal over $t_{MA} = 100$ ps. Reproduced from [320].

5-to-1 layout, with PRL-MN nodes receiving electronic signals in the main layer, and with an optically triggered PRL-MN node at the output layer. The presynaptic (in relation to the interlink) nodes in the main layer are all biased in the valley in (PDR II) operation point of the device, $V_0 = 770$ mV. Here, all the nodes are operated at

the same biasing point (and therefore have the same distance from their excitable threshold), but this is not a requirement for operation of the system, with further bias tuning of individual nodes allowing for increase or decrease in node sensitivity to inputs. The postsynaptic (output) PRL-MN node is also biased in a valley operation point (in PDR II). A temporally multiplexed sequence of varying spatiotemporal pulse patterns serves as input feed to the network. Each spatiotemporal pattern consists of up to five super-threshold pulses (up to one per each parallel input) of length $t_{pulse} = 80$ ps. These spatiotemporal patterns are fed into the main layer of neurons, where they elicit guaranteed spiking responses. These optical spiking signals are weighted through optical power attenuating elements (represented as a multiplication of the signal trace by a given factor w_n in the numerical model). Furthermore, to enable a labelled (supervised) learning approach in the model, an additional channel with guiding binary signal carries the data labels for target and non-target patterns, encoded in its amplitude (voltage). The output postsynaptic node performs signal summation (integration), thresholding and spike firing. A schematic diagram of the network is depicted in Fig. 5.3, demonstrating how different spatiotemporal pulse patterns (in blue) may elicit spiking responses and how a network with a particular weight setting results in an output spike activation only for a selected weighted pattern that sums up as sufficiently powerful postsynaptic signal. The pattern recognition procedure using the spiking network is shown in Fig. 5.3(b), using multiple time-multiplexed patterns with a temporal length of 420 ps per sample, representing a node processing capacity (rate) of up to 2.38×10^9 spikes/sec, or a full network processing capacity of 11.9 Gbps.

5.2.1 Hybrid supervised-STDP learning in the RTD-SpikeSLP

The design of individual building blocks and the architecture of the ANN represent parts of the overall challenge of developing the system. Another important part to tackle is the development of effective network training algorithms. A brief, general overview of SNN training approaches is provided in Sec. 1.2.2.

In accordance with conventional neural networks, the RTD-SpikeSLP is operated in two distinct phases: a) *training* phase and b) *inference* phase. To train the RTD-

SpikeSLP network model, an offline supervised learning rule is introduced following the approach described in [321] for training memristor-based physical neural networks. The regime of operation differs from the memristor-network, since the RTD-SpikeSLP is a SNN model that computes and propagates information using spike trains in the optical domain. This allows both for increased encoding sparsity, as well as to seize some of the general advantages of optical signalling (low loss waveguiding, no resistive heating, non-interacting nature of the signals).

First, in the training phase, the state of the network is adjusted via a dedicated (learning algorithm) to properly configure the system for inference. In this step, labelled patterns are fed into the network, and the output state is compared with each pattern's label provided in the additional teacher signal channel. Based on the output, adjustments are made according to the network weight matrix. This learning phase is performed over multiple repetitions (called epochs). In each epoch, the dynamics of each optoelectronic RTD-MN neuron within the system are evaluated. The use of a teacher signal allows for batching (processing of multiple patterns in a single epoch) by means of time-domain multiplexing per single epoch. The training procedure is successfully terminated once the weights stabilise. In the current iteration of the model, this is done by the user (operator), with an automated rule for decision making on reaching the optimal network (weight) state being considered for future versions of the system. Furthermore, as the weighted signals are received by the photodetector of the PRL-MN node (which directly converts the signals to current using a given factor κ), the absolute weight values depend on this current conversion (sensitivity) factor κ of the PD term. In this model, κ was slightly adjusted to set the bounds for the weights (multiplicative signal factors) within the usual float-valued interval $[0,1]$.

The evolution of the weight matrix during the training procedure is shown in Fig. 5.4. The weight matrix is initialized with each weight set to $w = 0.4$ and two different cases of target patterns are tested: $[0 \ 1 \ 0 \ 1 \ 0]$ (Fig. 5.4(a)) and $[1 \ 0 \ 1 \ 0 \ 1]$ (Fig. 5.4(b)). In each epoch ($t_{\text{epoch}} = 5 \text{ ns}$) of the learning phase, three temporally-multiplexed independent 5-bit spatial spatiotemporal patterns are picked, with a probability $P_t = 0.25$ of picking the target and $P_f = 0.75$ of picking any other pattern. This triplet,

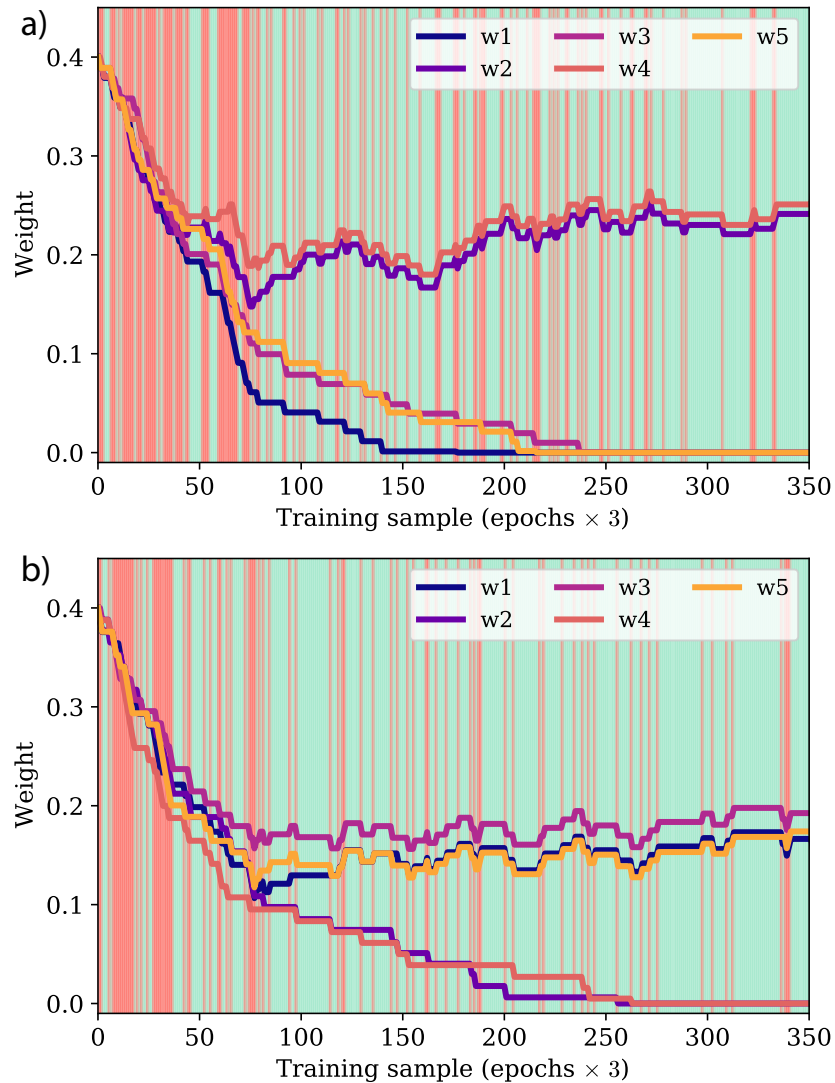


Figure 5.4: Demonstration of the supervised learning process for the spatio-temporal patterns. *a)* $[0\ 1\ 0\ 1\ 0]$ and *b)* $[1\ 0\ 1\ 0\ 1]$. Randomly sampled labelled patterns are used for the training via the custom learning rule. The background colour shows the network state (True/False) during each step. Reproduced from [320].

accompanied by the corresponding teacher signals, is then processed by the network.

The outcomes for each pattern are shown as the background highlighting colour in Fig. 5.4, where green represents *True positive*, *True negative* outcomes while red represents *False positive*, *False negative* outcomes. For either of the two *True* output states, no weights are adjusted during the learning step as the network has returned the desirable result (it has recognized the pattern). In the case of the *False positive*

output state, the weights that contributed to the firing are de-potentiated (weakened), with Δw representing a function capturing the presynaptic-postsynaptic spike temporal separation. As the temporal proximity between two spikes can be interpreted as a direct measure of their mutual influence, the closer in time the presynaptic spike is to activation of a *False positive* postsynaptic spike, the more pronounced is the applied depotentiation (weight weakening) effect. This can be considered as a supervised variation on the general spike-timing dependent plasticity (STDP) learning protocol. For Δw , a simple rational function was selected [320] for the weight adjustment:

$$\Delta w_n = \frac{a}{b \cdot |\Delta T_n| + c} + d \quad (5.2)$$

where

$$\Delta T_n = T_{POST} - T_{PRE,n} \quad (5.3)$$

represents the temporal separation (time interval) between the postsynaptic and presynaptic neuron n , $a = 9.35 \cdot 10^{-3}$, $b = 5 \cdot 10^9$, $c = 0.8$, $d = 1.5 \cdot 10^{-3}$. The numerical coefficients in the rational function were selected based on an observed typical temporal distances between spikes in presynaptic and postsynaptic neurons, and the corresponding desired weight adjustments. The weighting curve formulated by Eq. 5.2 as a function of spike separation time can be seen in Fig. 5.5.

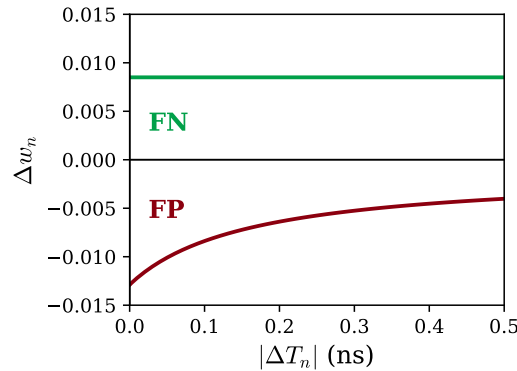


Figure 5.5: Weight adjustment factor Δw_n as a function of POST-PRE spiking interval $|\Delta T_n|$. For false negatives (FN, in green), the weight adjustment is a constant fixed value. For false positives, the weight adjustment magnitude is a function of $|\Delta T_n|$, with closer spikes yielding stronger depotentiation. Reproduced from [320].

5.2.2 Inference in the RTD-SpikeSLP model

With the presented model settings, the randomized sampling criteria and the provided hyperparameters, the occurrence rate of False (unwanted) outcomes during training gradually decreases until a stable weight matrix state is reached in approximately 300 tested patterns (which corresponds to 100 training epochs). The model only allows for positive weight values w_i (following the bounding rule set above) to ensure operation within the domain of physical feasibility (as attenuation cannot be negative and negative light intensity does not exist). After the training phase, the network can perform inference for recognition of the selected spatiotemporal 5-bit pattern. During inference, patterns with equal number of active bits were tested against the single desired target pattern: $[0\ 1\ 0\ 1\ 0]$ in one iteration, and $[1\ 0\ 1\ 0\ 1]$ in the second iteration. The results for inference and its accuracy for both of these cases are shown in the confusion matrices in Fig. 5.6.

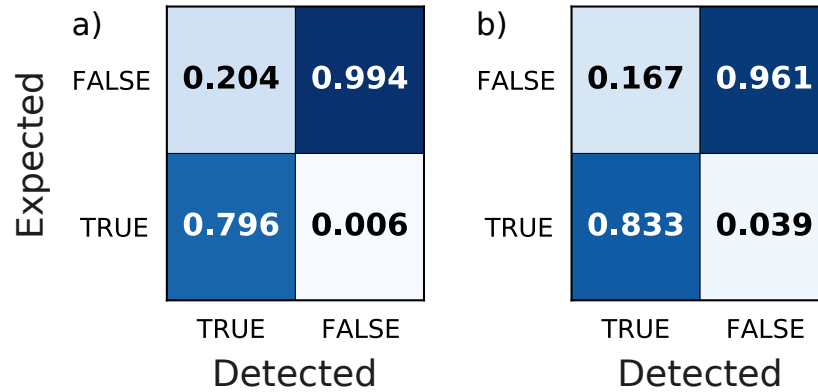


Figure 5.6: Confusion matrices for *a)* for inference of pattern $[0\ 1\ 0\ 1\ 0]$ against all other patterns with two ON bits ($n = 10$ different patterns, 540 total inference steps); *b)* for inference of pattern $[1\ 0\ 1\ 0\ 1]$ against all other patterns with three ON bits ($n = 10$ different patterns, 540 total inference steps).

In case of $[0\ 1\ 0\ 1\ 0]$ representing the target pattern, the total space of 5-bit patterns with $n_{ON} = 2$ active bits equals $\binom{5}{2} = 10$. In total, 540 inference steps were utilised, which means the whole set of available patterns was repeated $54\times$ during inference. In this case, the total True response accuracy was 97.4%. In the second case of $[1\ 0\ 1\ 0\ 1]$ representing the target pattern, the total space of 5-bit patterns with

$n_{ON} = 3$ active bits equals $\binom{5}{3} = 10$ as well. In total, 540 inference steps were utilised, which means the whole set of available patterns was repeated $54\times$. In this case, the total True response accuracy was 94.8%. Therefore, the trained model of the physical RTD-MN powered spiking single layer perceptron clearly demonstrates 94%+ accuracy for recognition of specific spatiotemporal pulse patterns.

5.2.3 Discussion of the RTD-SpikeSLP model

The results in this Section validate the operation of optoelectronic perceptron based upon the physical RTD models, and demonstrate its capability to process spatiotemporal data. To further improve the functionality of the model and range of potential data inputs it can process, there are additional properties that warrant further research interest. Currently, only positive weighting factors can be implemented on the optical feed-forward pathways (as light cannot have negative intensity). This currently limits the performance of the RTD-MN based perceptron to processing of similar patterns only. A possible solution for this could be devised by instead operating neurons in pairs, relying on RTDs unique capability of spiking in both peak and valley with mutually inverted responses. In such "balanced" layout, one RTD could provide the excitatory signals to the laser, while the second one could operate with the laser biased above the lasing threshold and utilise the excitability-induced drops in current I to provide inhibitory impulses. The weighting function would utilise switching between these two pathways to realise a full range of excitatory/inhibitory responses. However, the cost of such approach lies in both increased spatial and interlinking footprint as well as in the increased power consumption, particularly with the second additional balanced node operated with the laser biased above the lasing threshold.

Second, the current learning scheme allows for operation of up to N -to- N networks comprising two layers. This is sufficient to validate the proof-of-concept, yet for the desirable utilisation in deep neural network arrangements, the training scheme requires further extension similar to back-propagation to correctly estimate the error across multiple layers.

Finally, translation of the theoretical model into an experimental form remains yet

to be demonstrated. Ongoing research efforts are focused on realisation of nanoscale RTD systems with incorporated laser [282] and photo-detecting elements, and various approaches are under investigation for realisation of weightable optical interlinking between multiple on-chip neurons [322].

5.3 Experimental analysis of information processing in a PRL node

5.3.1 Coincidence detection in a single PRL node

Coincidences and synchrony of signals are important phenomena used in the brain during information processing [323] as well as in the sensory pathways, including e.g. the auditory cortex [324]. As an example, in visual perception, coincidence detection functionality enables mirror symmetry detection [325]. Here, we experimentally

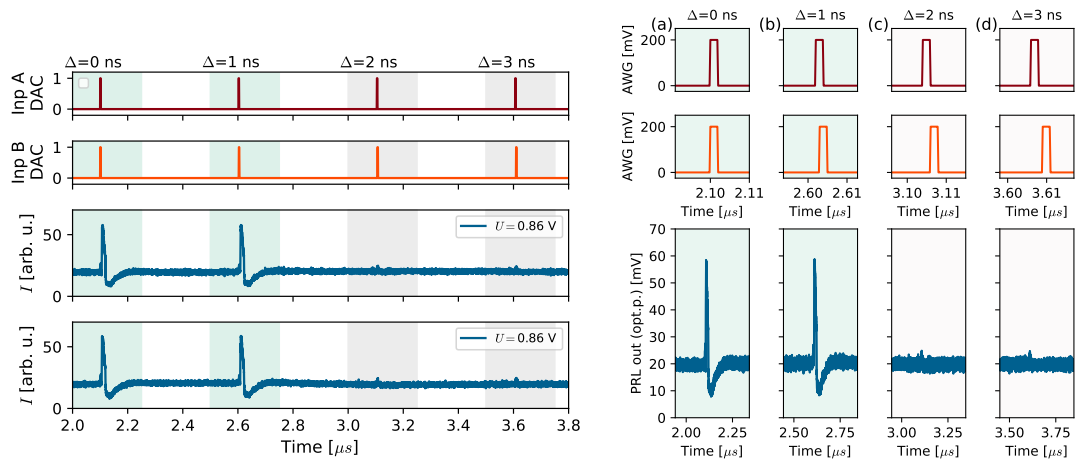


Figure 5.7: Two-input coincidence detection (logical AND) realised in the PRL node. In both left and right figures, the first two rows show the 2 ns square pulses (input stimuli) being gradually mutually shifted in time, in 4 different shifts of 0 ns to 3 ns. For perfect (leftmost) and for partial (2nd from the left) overlap, the total energy at a given moment surpasses the firing threshold, eliciting a spike. However, as the pulses drift further apart (c,d), the input perturbation energy to the system at any given time is not sufficient for spiking. Reproduced from [267].

demonstrate a coincidence detection task performed on a pair of independent optical

inputs entering into a single PRL node, using the modular experimental system described and characterised in Sec. 4.4. Both of the input branches were modulated using 2 ns square pulses with gradually increasing mutual temporal separations. The separation is measured between rising edges of these trigger pulses, and the gradual addition increment is $\Delta_{incr} = 1$ ns. As can be seen in Fig. 5.7(a,b), only perfectly overlapping pulses $\Delta_{incr} = 0$ ns and partially overlapping pulses $\Delta_{incr} = 1$ ns result on the total sum of optical input power exceeding the excitability threshold, therefore resulting in an activation event of a single excitable spike in the PRL. For larger temporal separation values (such as $\Delta \geq 2$ ns), no responses are triggered. This functionality can be considered as a temporal version of the *AND* logical gate.

5.3.2 Exclusive OR (XOR) task in a single PRL node

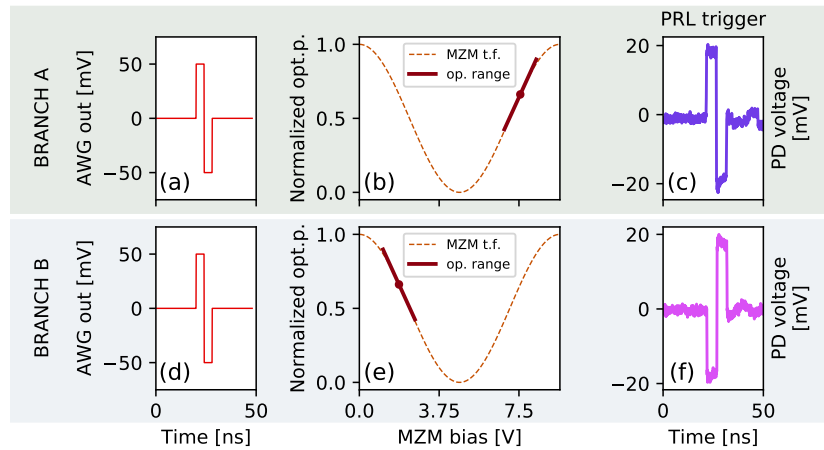


Figure 5.8: The pre-processing layer for the XOR task using the PRL, realised with MZMs. (a,d) The bipolar trigger pulses are processed through a layer of two (b,e) MZMs with different settings, providing mutually inverted input perturbations (c,f) on two optical input branches (denoted as A,B). Reproduced from [267].

The exclusive OR (XOR) task can be considered as one of the basic, two-input logic tasks. In an XOR gate, each individual input coming from one of the two branches results in a *TRUE* state, while both pulses coming together result in a *FALSE* state in the system. This "cancellation effect" makes the XOR task linearly inseparable and allows for practical "information subtraction". As such, XOR represents a fundamental

functionality for controlled propagation of information in larger neural networks - without cancellation, information (inputs) would only accumulate and the system would, in theory, lack the capability of processing linearly inseparable tasks. This used to be the case of single-layer perceptron networks [5] prior to the addition of extra layers. In relation to photonics, previously published works have demonstrated certain logic gates (AND, OR) realised in chains of spiking micropillar lasers with integrated saturable absorber [326]. This work now demonstrates the realisation of an XOR spiking gate using a single PRL node.

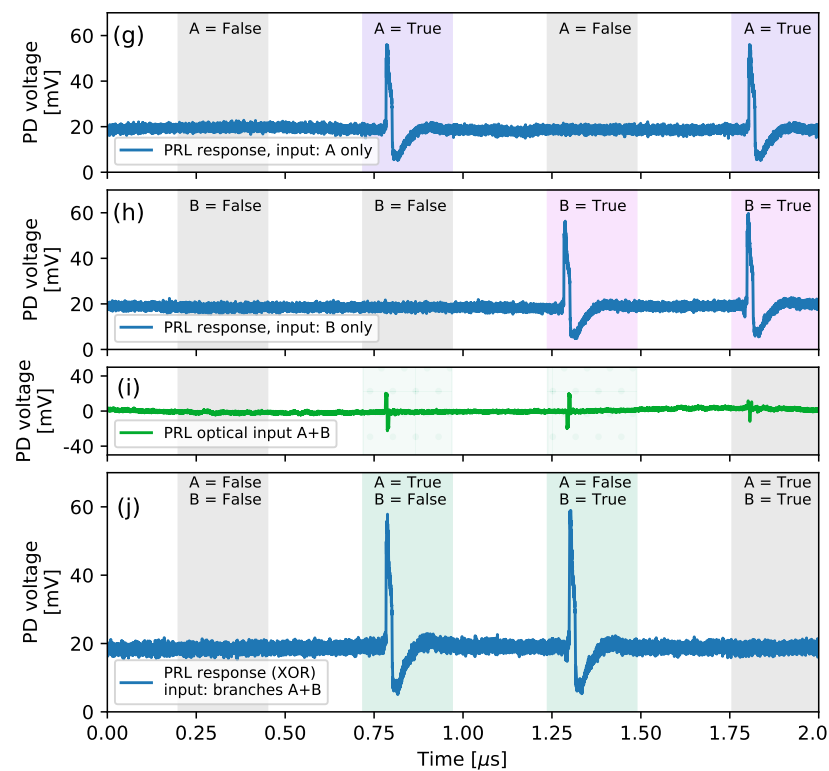


Figure 5.9: The exclusive OR (XOR) task demonstrated experimentally in the PRL. The input pulses (see Fig. 5.8) can individually elicit spiking responses from the PRL node (g,h), but when arriving simultaneously, they cancel out (i) and no spike is fired (j). This demonstrates the XOR task operation. Reproduced from [267].

The experimental results are collected in Fig. 5.9. Since an individual PRL spiking neuron is functionally closely related to the single-layer perceptron, it requires an additional (second) network layer which is here experimentally realised using two MZMs.

Each of the two input branches is fed with a signal from a tuneable laser, and provided with equivalent, bipolar, spike-like modulation signals (shown as red trace, 8 ns per bipolar pulse, Fig. 5.8(a,d)). The modulators convert these RF signals into optical modulation (Fig. 5.8(b,e)). For the XOR functionality, each of the MZMs is set in an opposing point of the modulator transfer function (at a point between quadrature and maximum), resulting in polarity inversion for the pulses in one of the branches (here, in branch B, Fig. 5.8(f)).

Following this "pre-processing" step via the layer of MZMs, the signals are then propagated to the PD element in the PRL node. There, each pulse individually is capable of eliciting an excitable spike thanks to the positive part of this perturbation exceeding the spike activation threshold (Fig. 5.9(g),(h)). However, for cases where the two pulses arrive simultaneously, they cancel each other out (Fig. 5.9(i)) as they are mutually in anti-phase. Therefore, no spike is elicited when A and B inputs are simultaneously active (Fig. 5.9(j)) and the full set of required states for the XOR task is satisfied.

In this experiment, the layer of modulators performed the operation of input pulse polarity inversion. However, it is important to point out that RTDs possess the unique capability of spiking with both polarities. This is allowed by the presence of the two operation points [300] (peak and valley biasing cases, which were demonstrated for example in Fig. 4.25). Therefore, the RTD itself can directly replace the function of the modulator in this layout and achieve the same functionality. This opens the way to RTD-based cascadable networks that enable handling of linearly inseparable data.

5.4 Dynamically-weighted optical interlink of RTD spiking nodes

Previous sections of this chapter have demonstrated some of the information processing capabilities of spiking RTD devices by the means of numerical modelling. In this final section of the thesis, a path towards experimental implementation of networked spiking optoelectronic RTD nodes will be investigated. A key property for the realisation of

spiking networked arrangements is *cascadability*. In cascable systems, a response from an upstream node can function as trigger for a downstream node. This is a fundamental functional requirement as it allows us to maintain and propagate signals during processing by regenerating them at sufficient (higher than decay) rates, suppressing noise and ensuring signal amplitude between processing layers being greater than unity [327]. In the realm of photonic devices, cascable responses have been shown in silicon-on-insulator microrings [180], micropillar lasers [256], excitable graphene-fibre lasers [328], integrated DFB lasers [329], Q-switched lasers [330] and VCSEL-SAs [234].

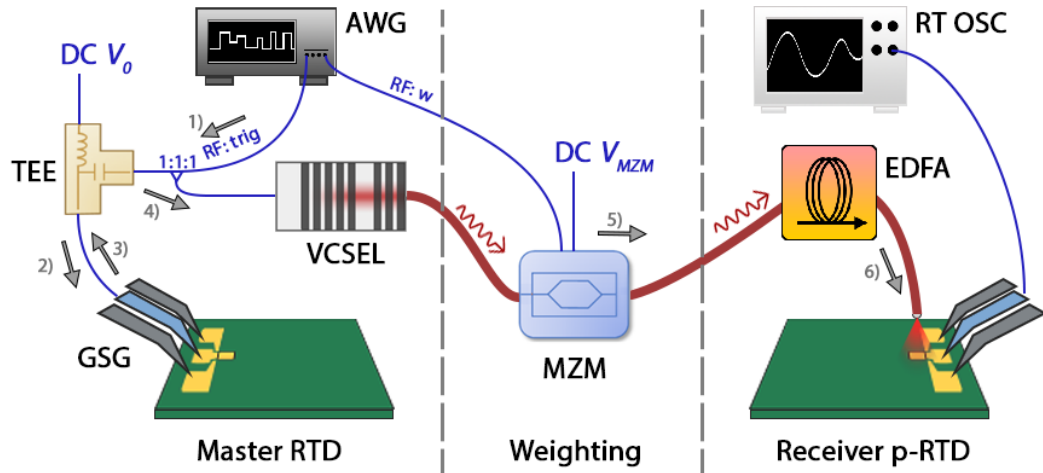


Figure 5.10: Scheme of the experimental setup and layout of the RTD-laser to PD-RTD interlink (or RL-PR in short) system with dynamical (MZM-enabled) weighting. Signals from one arbitrary waveform generator (AWG) channel provide electrical perturbations that activate electrical spiking in the master RTD. These spikes are used as modulation for a VCSEL, providing optical spiking regimes. These optical spikes are dynamically weighted at the MZM using second AWG signal. Finally, weighted spikes are directly injected into the receiver pRTD via a lens-ended fibre.

In this section, an experimental demonstration of cascability is provided by propagating spiking optical signals between two independently operated RTD systems. Furthermore, an additional MZM element is added to the path of the link, to enable for real-time, dynamical adjustment of amplitude of the propagated spike.

This experiment was performed by interlinking two different classes of devices using fibre optic components. The master μ -RTD-VCSEL node was realised using a single, high PVCr, $r = 3 \mu\text{m}$ diameter RTD fabricated at TU Eindhoven (TU/e; for I - V

and additional details regarding this device and I - V characteristic, see Sec. 4.4). The receiver was realised using a 700 nm mesa diameter, 14 μm optical window single pRTD device grown at the University of Glasgow. The benefit of using these devices in their respective roles is in the very high signal-to-noise ratio of excitable dynamics provided by the μ -RTD provided by TU/e and the light-absorptive capability of the pRTD fabricated by our collaborators at the Univ. of Glasgow. In this experiment, the master RTD is activated using electronic perturbations, and the receiver RTD is not coupled to a light source. Therefore, this demonstration can be considered as an E/O/E type of feed-forward interlink.

The experimental setup is shown in Fig. 5.10. An arbitrary waveform generator (AWG) provides two RF signals. One RF channel consists of series of square-shaped pulses, which provide spike trigger perturbations (stimuli) for the master RTD. A second RF channel from the AWG provides the dynamical weighting envelope. The electrical spiking signals from the master RTD are converted to optical spikes via direct modulation of a 1550 nm VCSEL (RF input is AC-coupled directly to the laser through a bias-tee). The VCSEL was operated with a bias current of 2.2 mA, slightly above its lasing threshold of approx. 1.9 mA at room temperature ($T = 298$ K). Additional details of the components used in this demonstration can be found in the sections describing the PRL node (Sec. 4.4) and the nano p-RTD (Sec. 4.5).

The resulting time traces of the experiments are shown in Fig. 5.11. To obtain excitable spiking and optical output functionality, the master node was forward-biased with $V_0 = 830$ mV (biasing in the peak region of the I - V). The spiking regimes in the master μ -RTD were triggered by an RF modulation signal consisting of 500 ps square shaped pulses with amplitude $V_{\text{amp}} = 75$ mV and a temporal separation of 250 ns between pulse rising edges (Fig. 5.11(a)). The pulse polarity was matched to the forward-biased, peak operation point biased RTD, with pulses being positive in voltage. The activation of excitable spikes using electrical signals provides a high degree of activation reliability.

Once electrical spiking was elicited in the master RTD, these signals were then E/O converted into optical spikes by the VCSEL. The latter's optical output (Fig. 5.11(b))

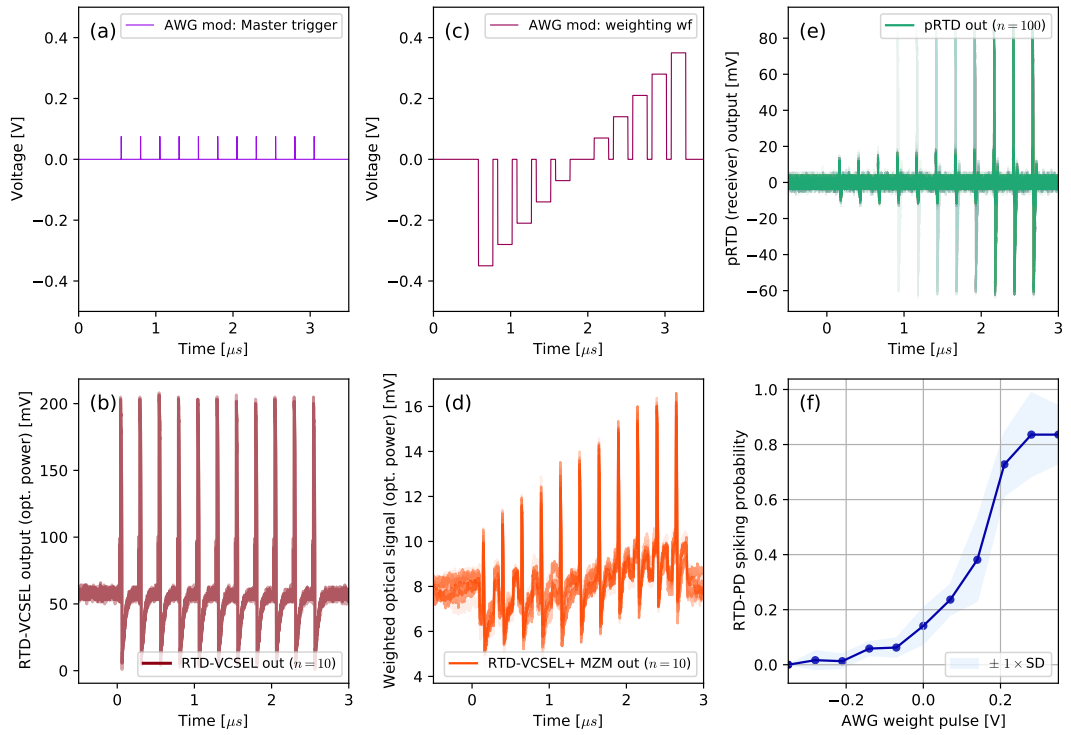


Figure 5.11: Demonstration of dynamic pulse weighting in a master-receiver layout of two RTD optoelectronic nodes (PRL \rightarrow pRTD arrangement). (a) The electrical trigger pulses (stimuli) used to elicit excitable spiking in the master μ -RTD. (b) Optical spiking signals from the master RTD (biased in peak region) coupled to a VCSEL (for E/O conversion), in response to the stimuli. (c) Dynamical weighting waveform applied to the synaptic MZM element on the interlink. (d) Dynamically weighted optical spikes from the master RTD (RTD-VCSEL). (e) Responses of the receiver (p-RTD) to the weighted optical spikes, injected directly into the p-RTD via a lens-ended fibre. (f) Histogram of spike elicitation probabilities as a function of spike weighting factor (spike amplitude).

represents the input of the interlink between the two RTDs. To realise the interlink, the VCSEL's optical signal with the optical spikes (with CW-average power of $\approx 200 \mu\text{W}$) was directed to an erbium-doped fibre amplifier (EDFA, Thorlabs EDFA100S) to boost its level. This was followed by a fibre polarisation controller (PC, to adjust polarisation prior to the modulator) and a MZM (Thorlabs) for weighting. The last two components are optional: while these are not strictly required to realise the basic interlinking functionality, they allow to perform dynamical signal modulation for such use-cases as weighting of the communicating optical spikes from the master to the receiving RTD

nodes. Additionally, these MZMs also offer an option to provide additional signals for the downstream node.

To perform the weighting functionality, the off-the-shelf MZM (Thorlabs) was at a fixed operational point (voltage) in the semi-linear region of the modulator transfer function (the same as in previous cases, see Fig. 5.9(b,e)), and modulated with an RF signal consisting of gradually varying square-shaped pulses with return-to-zero (RZ) characteristics. This RF "staircase" signal is shown in Fig. 5.11(c). These weighting steps are 187.5 ns in length, interleaved with return-to-zero sections of 62.5 ns (corresponding to 250 ns pulses with 75% duty cycle). Because the MZM operates as a light intensity modulator, the optical spikes are amplitude modulated. Because the VCSEL is operated very close to its lasing threshold, there is minimal CW output (lasing pedestal), resulting in mostly the amplitude of spikes being adjusted. The dual polarity character of the modulation signal was used to better utilise the maximum available peak-to-peak RF signal amplitude generated from the digital-to-analog converted (DAC) of the AWG. The resulting weighted optical spikes are shown in Fig. 5.11(d), demonstrating the well controlled character of the weighting operation in the synaptic optical interlink.

Finally, the response of the receiver p-RTD to the input of the weighted optical spikes (Fig. 5.11(d)) is shown in Fig. 5.11(e). The p-RTD is biased at $V_0 = 745$ mV (biasing in the valley region of the I - V , shown in Fig. 4.21). The input to the p-RTD was realised via amplification of the signal with weighted optical spikes using an EDFA, and then through coupling with a lens-ended fibre (OZ Optics) into the exposed optical window of the p-RTD (target average optical power after amplification of 250 μ W). To allow for better statistical evaluation of the produced dynamical responses in the receiver p-RTD, $n = 100$ subsequent readouts were acquired from the real-time oscilloscope. These are temporally aligned and overlaid in Fig. 5.11(e), resulting in higher plotting contrast for more frequently occurring events (spikes). This visualisation shows and confirms that weighted spikes with lower absolute amplitude directly correspond to lower spike elicitation probability in the receiver p-RTD. Furthermore, the number of occurrences for each individual weighted spike was counted over all the recorded readouts, and plotted as a histogram in Fig. 5.11(f). This further validates the observation that

optical spike triggering in the receiver RTD directly corresponds to the dynamically adjusted weighting value applied on the optical spikes in the synaptic interlink. This demonstration represents the first reported case of direct optical spike propagation between two optoelectronic RTDs. The possibility of cascable triggering of excitable responses in the downstream optoelectronic node solely using optical signals carrying RTD-produced spikes represents an important step towards extended interconnected layouts and RTD-enabled optoelectronic spiking networks.

Chapter 6

Conclusions and Outlook

6.1 Thesis summary

THE systems investigated in this thesis can be classified as belonging to two main directions: (1) all-optical VCSEL-based spiking neurons, and (2) optoelectronic RTD-based spiking neurons. Each of these systems was studied both in experiments (primarily), and also in theory with numerical models. This is highlighted in a generalised overview diagram shown in Fig. 6.1.

6.1.1 All-optical artificial spiking VCSEL-neurons

Building upon the previous demonstrations and published works on excitable spiking phenomena in telecom wavelength VCSELs, this work has expanded the domain of experimental photonic VCSEL-neurons based on injection locking in three main directions. First, advanced characterisation of the excitable phenomena was conducted in experiments, including study of the characteristics of nonlinear dynamical behaviours of interest, such as spike latency onset delay induced by varying input perturbation amplitudes [255]. Second, the concept of spiking rate-coding was introduced and experimentally studied in the spiking VCSEL-neuron. This behaviour was utilised to perform encoding of digital images into a fully temporal, spike based representation, demonstrating the viability of using VCSEL-neurons as all-optical, high speed encoders

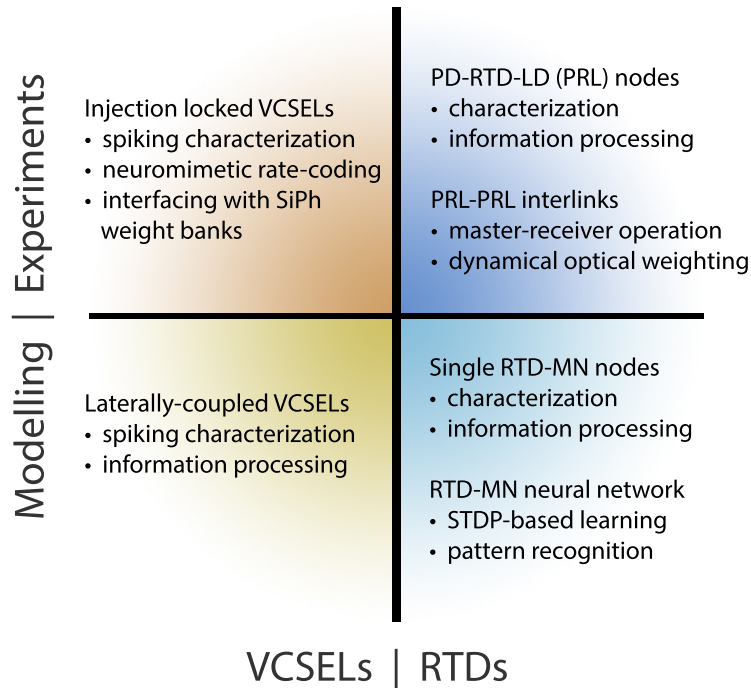


Figure 6.1: Overview of thesis results.

for data conversion in the temporal domain, as utilised in neuromorphic engineering [254]. Third, the combined operation of VCSEL-neurons in conjunction with silicon photonics (SiPh) based integrated microring weight banks was experimentally demonstrated. This was performed using two fundamental connected layouts: (1) amplitude weighting of the ultrafast optical spiking signals produced by the VCSEL-neuron using a microring of the SiPh chip (representing a synaptic functionality); and (2) performing rate-coding on signals that were directly intensity modulated in the microring (representing a neuromimetic temporal encoding functionality with signal directly from the SiPh platform). This work was performed in collaboration with Prof Paul Prucnal’s Lightwave group at Princeton University (NJ, USA).

Furthermore, in a collaboration with the team of Prof Mike J. Adams at the University of Essex, a numerical study was performed on the prospects of laterally (evanescently) coupled VCSELs for operation as spiking optical elements in neuromorphic systems [331]. Laterally-coupled VCSELs offer an alternative system that is by-design fully chip-scale,

and can also exhibit neuron-like spiking regimes. These have been explored numerically, successfully demonstrating dynamical phenomena such as excitable and rebound spiking, continuous firing rate coding and digital-to-spike datastream conversion. Full details of this work can be found in a recently published manuscript [331].

In summary, these experimental and numerical studies were aimed at expanding our understanding of photonic VCSEL-neurons, with practical brain-inspired information processing applications as a main aim. In total, three studies were first-authored and published in peer reviewed journals [254, 255, 331], with the VCSEL-MRR study currently available as a pre-print [264] and under peer review. In particular, the rate-coding study [254] and the evanescently-coupled VCSEL study [331] demonstrated and evaluated new spike-based dynamical phenomena in the VCSEL-neuron, further strengthening its position as a prime candidate for realisation of spiking neurons in neuromorphic photonic systems.

6.1.2 Optoelectronic artificial RTD-neurons

The second class of studied systems was optoelectronic circuits based on resonant tunnelling diodes (RTDs). Unlike the case of VCSELs, RTDs have not been prominently featured in the domain of neuromorphic optical computing until 2020, with just a few prior studies [289, 300, 332] making the connection between the observable dynamics in RTD optoelectronic oscillators and neuromorphic functionalities. This has been significantly expanded by the research performed within the EU Horizon2020 project 'ChipAI', under which the research shown in this thesis was conducted.

As introduced in Sec. 4.1, the RTD is an active semiconductor mesoscopic device. Its unique properties stem from an embedded double-barrier quantum well (DBQW) within the heterostructure, which allows for quantum tunnelling of electrons through the discrete (resonant) state within the DBQW under specific external bias conditions. Thanks to this, the RTD exhibits a highly nonlinear voltage-current (I - V) characteristic, as well as the ability to operate with extremely high bandwidths thanks to the ultrafast timescales of quantum tunnelling through the DBQW. As a result, RTDs exhibit nonlinear dynamical responses. These typically include self-sustained oscillations, but

also extend to more exotic dynamics such as excitability. Excitable RTDs form the main focus of Chapters 4 and 5.

The RTDs were studied both experimentally and using numerical models. First, an overview of the RTD characteristics and properties was introduced, and a dynamical model based on ordinary differential equations was discussed. Since the RTD is an electronic device, an extension toward optoelectronic operation regime is proposed. Finally, a fully opto-electro-optical (O/E/O) excitable spiking artificial neural node based on an RTD was proposed in theory and demonstrated experimentally.

In experiments, a first demonstration of the O/E/O RTD powered excitable RTD node was reported in [267]. This was realised by RF coupling a micrometric RTD (fabricated on an InP platform by research collaborators at TU/e, The Netherlands) to an active (amplified) photodetector and a VCSEL (through a bias-tee network). It was shown that the RTD can produce well defined, high SNR excitable spikes with a clearly defined excitability threshold; that these spiking regimes can be activated deterministically by optical trigger signals (perturbations) received by the coupled photodetector, and that the electrically modulated, RF-coupled VCSEL converts these electrical spikes into the optical domain. Furthermore, the refractory period of the system was studied and evaluated, and two information processing tasks (coincidence detection and XOR logic operation) demonstrated for the case of two (optical) input signal branches arriving simultaneously at the photodetector. These tasks rely on the excitability functionality of the RTD, and include a coincidence detection (which can be considered functionally equivalent to an AND logical gate), and an exclusive OR (XOR) task realised in the RTD with an additional "virtual" processing network layer realised physically using the MZMs on each input branch. In both cases, the spiking logical gates were successfully experimentally demonstrated, and an additional validation for the coincidence detection task was provided using the numerical model, showing the system's potential for operation at GHz rates.

In addition, an alternative experimental setup was reported based on an RTD that directly embeds an optical (photo-absorptive) window within the heterostructure, allowing the device to respond to input light signals without the requirement for a

coupled external photodetector. This so called p-RTD (photodetecting RTD) was shown to also exhibit deterministically triggered excitable spiking regimes, clear spike activation threshold, and refractory times when biased both in the peak and the valley regions of the I - V characteristic. Furthermore, operation was demonstrated for a p-RTD coupled to a VCSEL, again realising a full O/E/O node with a reduced number of required external components and a simplified coupling circuit. Additionally, a novel resonate-and-fire spiking behaviour was reported for the first time, further highlighting the potential of RTDs for realisation of optoelectronic excitable network nodes for use in unconventional, light-enabled, brain-inspired computing. Finally, a direct in-device pre-synaptic spike weighting functionality was reported in the coupled RTD-VCSEL system [333].

Numerically, the prospects of nanoscale, RTD-powered O/E/O nodes for more complex information processing functionalities were also investigated in [320]. First, this included serial data processing of PAM-pulse sequences using the short-term temporal memory of the dynamical system. Using this approach, a pattern recognition task was realised in the excitable nanoRTD node for serialised, binary Tetris-like patterns. Furthermore, a feed-forward single layer spiking perceptron (SpikeSLP) neural network architecture was proposed, utilising the optoelectronic nanoRTDs as network nodes. Such a network is powered by the nonlinear dynamical RTD models as spiking neurons, and allows for asynchronous, fully temporal computation. The information processing capability of this network was demonstrated on an input dataset consisting of spatiotemporal binary patterns. A custom learning rule for the network was proposed and utilised, combining the principles of brain-inspired Hebbian learning and supervised learning, and good degree of accuracy was demonstrated for spatiotemporal patterns with varying bit length.

6.2 Future outlooks

The RTD systems presented and investigated in this thesis represent an emergent technology that is currently not widely used in commercial settings. On the nine-level

scale of technology readiness levels (TRL) [334], RTDs can be considered as TRL 4, and neuromorphic system based upon RTDs can be considered as TRL 2. At the current development phase of these neuromorphic optoelectronic systems, there are multiple important steps ahead needed to increase the TRL. Perhaps as a first step before venturing deeper into the realms of novel architectures and unconventional computer hardware (be it for AI acceleration or brain-inspired computation), it will be important to find a common ground and a mutually understandable perspective among the various research directions within the field neuromorphic research. For example, focusing solely on the sub-field of photonics, photonic device designers need to understand the needs of circuit designers, and vice versa. The algorithm designers and electronic engineers need to be involved to ensure good hardware-software co-design and optimal functionality. We certainly need to keep collaborating with computer scientists and AI engineers to ensure any proposed hardware solutions are well matched to the needs of the industry. Finding common ground and understanding among so many disparate fields is no small feat, and will remain a crucial (ongoing) effort to unlock the field's full potential as we progress.

While the following list discusses some of the challenges of further development of the technologies studied in this thesis in a device-to-system order, they are not intended to be solved one after another in that given order. On the contrary, all of these questions remain open for solution in parallel, and will likely mutually influence and shape each other as new insights and better understanding arise.

6.2.1 Device level

First, the optimal fabrication procedure of miniaturised photonic components needs to be investigated to allow reliable, repeatable and high yield production of such devices. With recent rapid developments in the field of photonic integration and the emergence of photonic integrated circuits fabricated by foundries-as-a-service (such as Ligentec, JePPiX etc), new effective and more advanced solutions are likely to keep arriving at an increasing pace.

In the case of the RTDs, optimal structure design and epilayer parameters remain a

focus of ongoing research [281]. Furthermore, photodetecting RTDs need to be optimised for higher responsivity and lower spatial footprint. Previous works have demonstrated ultra-sensitive RTD photodetectors [274], highlighting the feasibility of such a solution. One of the advantages of the RTD technology is the ability to enable all the required functional O/E/O blocks of the neuromorphic optoelectronic nodes in a single platform. Furthermore, one of the interesting advantages of nanoscale RTD optoelectronic nodes may also come from their low footprint. Previous review articles [149] have stated that the minimal lateral size of artificial photonic neurons is expected to be around $100\mu\text{m}$. For the nano-RTD node, this is likely to be notably smaller.

Besides the RTD device itself, miniaturised and on-chip light sources operating in conjunction with the RTD need to be demonstrated. Here, an advantage of the RTD-powered networked arrangements is the flexibility of choosing either coherent or non-coherent light sources. Sub- λ RTD-LEDs can operate at high speeds with very low power consumption and offer a very good solution for shorter distance, low-coherence requirement interlinks [150]. From coherent sources, III/V integrated laser devices are currently the most promising. An interesting path forward is offered by nanowire-based lasers, as they are expected to have improved operation speed and reduced lasing threshold [335] thanks to their reduced size. Unfortunately, robust nanowire lasers are still to be delivered, with one larger scale statistical study demonstrating that material quality rather than cavity quality primarily influences the lasing operation [336]. Low photon count signalling with high light extraction efficiency was previously demonstrated by precisely fabricating active centers within nanowires [337].

6.2.2 Functional interlinking

Third, a feasible, chip-scale interlinking scheme needs to be found for these RTD-based optoelectronic networks. Here, a promising solution is offered by 3D microprinted waveguides fabricated using two-photon polymerisation (2PP). These have been demonstrated in the form of suspended, bridge-like waveguides [338] and fully three-dimensional meshes of interlinks with directly demonstrated fan-out (1-to-N) capability and CMOS compatibility [339, 340]. Among in-plane signal routing solutions, integrated photonic

routing manifolds [341] and integrated photorefractive-material based structures with embedded synaptic weighting functionality [322] are very promising solutions. The fan-in functionality of merging many-to-one waveguides for input into a given downstream node might be realised with integrated multimode interference couplers [342].

In addition, this includes the challenge of finding the most suitable technology that will govern the network state (weighting technology). Factors such as device state volatility and non-volatility, various weight update speeds and all-optical or optoelectronic signal modulation technology are offered by different photonic technologies and material platforms, with no clear winner currently in sight. A related challenge is in realisation of photonic memories, which are significantly more difficult to produce than their electronic counterparts. While it is likely that not every neuromorphic photonic system will have to be stateful, the question of memories remains important. Here, the ability of autaptic (recurrent) links of spiking neurons may be an interesting research direction.

6.2.3 Operational protocols

A fourth challenge lies in finding an operational protocol that will truly make the system stand-out in terms of performance. This challenge can be considered as shared among the diverse approaches for spike-based photonic hardware. This includes seizing the capabilities of temporal spike-based encoding for sparse and efficient information representation. The addition of the temporal domain into computation can be simultaneously considered as a significant advantage and a challenge. This is because our computers are digital and synchronous, and interfacing with asynchronous machines such as neuromorphic chips might be non-trivial. This also relates to the challenge of benchmarking: currently, the research community usually benchmarks performance of new neuromorphic architectures against conventional digital datasets such as the MNIST handwritten digits (which are in essence digital bitmaps), yet truly neuromorphic (domain specific) benchmarks that will be temporal, parallel or graph-based will likely be needed for proper performance evaluation. Notably, there are some recent advances in the direction of more unified benchmarking for neuromorphic hardware [343]. In a

similar way to that in which backpropagation became a golden standard for artificial neural network training in the field of deep learning, the emergence of a comparable "winning" approach will likely be required to achieve a critical mass for spiking neural networks. Furthermore, such a scheme must also be sufficiently hardware-feasible to unlock the prospective high performance of neuromorphic photonic platforms. Learning in photonic hardware was demonstrated using gradient-based approaches based on the finite difference method, using forward propagation [141] or backpropagation [159] in MZI meshes, in either case operating with continuous wave signals. Beyond these, alternative neural network approaches might also offer feasibility for hardware integration, including binarized networks [344] which have been shown to operate with good performance and efficiency [345] despite their binary-only synaptic links. Finally, newly emerging approaches for neural network learning such as analog implementation focused, local equilibrium propagation [346] and forward-forward propagation [347] may yield back-propagation alternatives that will be more hardware-friendly.

6.2.4 System-level approach

Finally, a neuromorphic system will (in the author's opinion) never operate fully independently from other digital systems. Rather, if neuromorphic platforms succeed in their goal to deliver effective AI-tailored computation, they will become a class of high-performance, application-specific accelerators augmenting our conventional digital computers. For this reason, it is important to consider these novel systems through the lens of a system approach. Realizing full-scale functional systems will require not only optical/photonic components, but also the accompanying high-frequency electronics, microcontrollers, FPGAs and interfaces. Such true system-level thinking can be seen for example in [138], and will only become more prominent as the field further matures.

The competition within the neuromorphic computation field is growing rapidly, and to remain competitive with other approaches it will be key to utilise photonics in a way that exploits its most powerful properties. Dedicated neuromorphic hardware will rely on elements for computing (nonlinearity, memory, gain), communication (with fan-in/fan-out), likely certain linear algebra operations (MAC, MVM) and the remaining

controllers. Electronics readily offer nonlinearity (transistor), easy memory (in the form of random access memories such as DRAM, SRAM) and easy access to gain (amplification). Electronic computing can be used to realise all the required building blocks of a digital computer, and therefore requires no domain crossing. In comparison, nonlinearity is comparably more difficult to obtain in optics, and effective photonic memories together with on-chip gain all remain a challenge. On the other hand, communications and fan-out in photonics are very efficient, with negligible cost per length of transmission (using waveguides, beam-splitters) while communications using electrons have a $\sim CV^2$ energy cost. Wavelength multiplexing remains a significant point of advantage for increasing the bandwidth of photonic systems. Furthermore, a very important challenge when realising such mixed-signal systems arises from the energy requirements for domain (A/D, D/A) conversions between photonic and electronic signals, and their relation to desired signal-to-noise ratios [348]. This will require careful design in order to minimise domain-hopping and increase energy efficiency. Avoiding D/A conversion at the system input (processing analog data) might be desirable to increase energy and computing efficiency.

In summary, the field of neuromorphic engineering lies at the convergence of hardware, applications and algorithms research. The use of light in neuromorphic systems provides access to low-loss data movement, very low latencies and high (GHz+) bandwidths. Together with the increasing importance of efficient computing hardware for mankind's ever-growing computational needs, and the challenges of further miniaturisation in digital electronics, unconventional chip architectures will keep moving more into the spotlight. Currently, the field of AI is focusing primarily on deep learning. However, despite the undeniable power and functionality of DL approaches, ultra-large DL models suffer from drawbacks such as extreme resource requirements and, in certain cases, unexplainability (a black box character). For these reasons, alternative AI approaches such as hyper-dimensional computing, graph neural networks and spiking neural networks are, in the author's opinion, likely to receive increased attention over the coming years.

Meanwhile, alongside the driving forces mentioned above, further developments in

the technology and fabrication in the field of photonic integration will ensure that the prospects of optical computing will grow stronger. Neuromorphic optical computing is currently still in the nascent development phase, without a clear-cut winner across the breadth of investigated approaches and architectures. Currently, a full-scale photonic AI accelerator with a clear advantage over state-of-the-art commercially-available electronics remains to be demonstrated. Nonetheless, the field is undergoing a significant expansion with a massive surge in interest, with a few startups already promising delivery of powerful photonic computing chips in the near future. In the author's opinion, thanks to all the previously mentioned factors coming together, the prospects of optical computing are as bright as they have ever been, and neuromorphic photonics represents an exciting direction of further research.

List of scientific contributions

First-authored publications

- ◆ **Hejda, M.**, Robertson, J., Bueno, J. & Hurtado, A. *Spike-based information encoding in vertical cavity surface emitting lasers for neuromorphic photonic systems*. Journal of Physics: Photonics 2, 044001. 10.1088/2515-7647/aba670 (2020) [255].
- ◆ **Hejda, M.**, Robertson, J., Bueno, J., Alanis, J. A. & Hurtado, A. *Neuromorphic encoding of image pixel data into rate-coded optical spike trains with a photonic VCSEL-neuron*. APL Photonics 6, 060802. 10.1515/10.1063/5.0048674 (2021) [254].
- ◆ **Hejda, M.**, Alanis, J.A., Ortega-Piwonka, I., Lourenço, J., Figueiredo, J., Javaloyes, J., Romeira, B. & Hurtado, A. *Resonant Tunneling Diode Nano-Optoelectronic Excitable Nodes for Neuromorphic Spike-Based Information Processing*. Phys. Rev. Applied 17, 024072. 10.1103/PhysRevApplied.17.024072 (2022) [320].
- ◆ **Hejda, M.**, Vaughan, M., Henning, I., Al-Seyab, R., Hurtado, A. & Adams, M. *Spiking Behaviour in Laterally-Coupled Pairs of VCSELs With Applications in Neuromorphic Photonics*. IEEE J. Select. Topics Quantum Electron. 1–10. 10.1109/JSTQE.2022.3218950 (2022) doi:10.1109/JSTQE.2022.3218950.
- ◆ **Hejda, M.**, Malysheva, E., Owen-Newns, D., Al-Taai, Q.R.A., Zhang, W., Ortega-Piwonka, I., Javaloyes, J., Wasige, E., Dolores-Calzadilla, V., Figueiredo, J.M.L., Romeira, B. & Hurtado, A. *Artificial optoelectronic spiking neuron based on*

a resonant tunnelling diode coupled to a vertical cavity surface emitting laser. Nanophotonics 0362. 10.1515/nanoph-2022-0362 (2022) [267].

- ◆ **Hejda, M.**, Doris, E.A., Bilodeau, S., Robertson, J., Owen-Newns, D., Shastri, B.J., Prucnal, P.R. & Hurtado, A. *Interfacing spiking VCSEL-neurons with silicon photonics weight banks towards integrated neuromorphic photonic systems.* arXiv:2305.00788 (2023) [264].

Co-authored publications

- ◇ Robertson, J., **Hejda, M.**, Bueno, J. & Hurtado, A. *Ultrafast optical integration and pattern classification for neuromorphic photonics based on spiking VCSEL neurons.* Scientific Reports 10, 6098. 10.1038/s41598-020-62945-5 (2020) [236].
- ◇ Robertson, J., Zhang, Y., **Hejda, M.**, Bueno, J., Xiang, S. & Hurtado, A. *Image edge detection with a photonic spiking VCSEL-neuron.* Optics Express 28, 37526. 10.1364/OE.408747 (2020) [238].
- ◇ Bueno, J., Robertson, J., **Hejda, M.** & Hurtado, A. *Comprehensive Performance Analysis of a VCSEL-Based Photonic Reservoir Computer.* IEEE Photon. Technol. Lett. 33, 920–923. 10.1109/LPT.2021.3075095 (2021) [27].
- ◇ Zhang, Y., Robertson, J., Xiang, S., **Hejda, M.**, Bueno, J. & Hurtado, A. *All-optical neuromorphic binary convolution with a spiking VCSEL neuron for image gradient magnitudes.* Photon. Res. 9, B201. 10.1364/PRJ.412141 (2021) [237].
- ◇ Alanis, J. A., Robertson, J., **Hejda, M.** & Hurtado, A. *Weight Adjustable Photonic Synapse by Non-Linear Gain in a Vertical Cavity Semiconductor Optical Amplifier.* Applied Physics Letters. 10.1063/5.0064374 (2021) [349].
- ◇ Robertson, J., Kirkland, P., Alanis, J.A., **Hejda, M.**, Bueno, J., Di Caterina, G. & Hurtado, A. *Ultrafast neuromorphic photonic image processing with a VCSEL neuron.* Sci Rep 12, 4874. 10.1038/s41598-022-08703-1 (2022) [239].

- ✧ Skalli, A., Robertson, J., Owen-Newns, D., **Hejda, M.**, Porte, X., Reitzenstein, S., Hurtado, A. & Brunner D. *Photonic neuromorphic computing using vertical cavity semiconductor lasers*. Opt. Mater. Express 12, 2395. 10.1364/OME.450926 (2022) [242].
- ✧ Ortega-Piwonka, I., **Hejda, M.**, Alanis, J.A, Lourenço, J., Hurtado, A., Figueiredo, J.M.L, Romeira, B. & Javaloyes, J. *Spike propagation in a nanolaser-based optoelectronic neuron*. Opt. Mater. Express. 10.1364/OME.451706 (2022) [313].
- ✧ Robertson, J., Alanis, J. A., **Hejda, M.** & Hurtado, A. *Photonic synaptic system for MAC operations by interconnected vertical cavity surface emitting lasers*. Opt. Mater. Express, OME 12, 1417–1426. 10.1364/OME.450923 (2022) [350].
- ✧ Zhang, W., **Hejda, M.**, Malysheva, E., Al-Taai, Q.R.A, Javaloyes, J., Wasige, E., Figueiredo, J.M.L., Calzadilla, V., Romeira, B. & Hurtado, A. *Tuneable presynaptic weighting in optoelectronic spiking neurons built with laser-coupled resonant tunneling diodes*. J. Phys. D: Appl. Phys. 56(8). 10.1088/1361-6463/aca914 (2022) [333].
- ✧ Haegele, S., Corrielli, G., **Hejda, M.**, Duempelmann, L., Terborg, R.A., Osellame, R. & Pruneri, V. *Large field-of-view holographic imager with ultra-high phase sensitivity using multi-angle illumination*. Optics and Lasers in Engineering 161, 107315. 10.1016/j.optlaseng.2022.107315 (2023) [351].
- ✧ Owen-Newns, D., Robertson, J., **Hejda, M.** & Hurtado, A. *GHz Rate Neuromorphic Photonic Spiking Neural Network With a Single Vertical-Cavity Surface-Emitting Laser (VCSEL)*. IEEE J. Select. Topics Quantum Electron. 29, 1–10. 10.1109/JSTQE.2022.3205716 (2023) [240].
- ✧ Al-Taai, Q.R.A, **Hejda, M.**, Zhang, W., Romeira, B., Figueiredo, J.M.L., Wasige, E. & Hurtado, A. *Optically-triggered deterministic spiking regimes in nanostructure resonant tunnelling diode-photodetectors*. arXiv:2304.11713 (2023) [319].
- ✧ Owen-Newns, D., Robertson, J., **Hejda, M.** & Hurtado, A. *Photonic Spiking Neural Networks with Highly Efficient Training Protocols for Ultrafast Neuromor-*

phic Computing Systems. SPJ Intelligent Computing. 10.34133/icomputing.0031 (2023) [352].

Oral conference contributions

- ❖ **Hejda, M.**, Robertson, J., Bueno, J. & Hurtado, A. *All-optical neuromorphic spike information encoding with VCSELs*. SIOE2020, Cardiff, UK. (05-03-2020). [Cancelled due to SARS-CoV-2 outbreak.]
- ❖ **Hejda, M.**, Robertson, J., Bueno, J. & Hurtado, A. *VCSELs for fast neuromorphic photonic systems operating at GHz rates*. Optical Fiber Communication Conference (OFC), San Diego, USA. 10.1364/OFC.2020.T4C.1 (12-03-2020). [Cancelled due to SARS-CoV-2 outbreak.]
- ❖ **Hejda, M.**, Robertson, J., Bueno, J. & Hurtado, A. *Neuromorphic systems based on photonics* Strathclyde Doctoral School Symposium (DSMS), Glasgow, UK. (26-05-2020). [remote]
- ❖ **Hejda, M.**, Robertson, J., Bueno, J., Alanis, J. A. & Hurtado, A. *Photonic VCSEL-neuron for spike-rate representation of digital image data..* CLEO/EQEC Europe, München, Germany. ISBN:978-1-66541-876-8 (24-06-2021). [remote]
- ❖ **Hejda, M.**, Alanis, J.A., Ortega-Piwonka, I., Lourenço, J., Figueiredo, J., Javaloyes, J., Romeira, B., Hurtado, A. *Photonic Spiking Neural Network with Resonant Tunneling Diode Optoelectronic Neurons*. CLEO, San José, USA. (18-05-2022).
- ❖ **Hejda, M.**, Owen-Newns, D., Zhang, W., Alanis, J.A., Bueno, J., Robertson, J. & Hurtado, A. *Photonic Neuromorphic Computing with Vertical Cavity Surface Emitting Lasers*. 5th International Conference on Applications of Optics and Photonics (AOP), Guimarães, Portugal. (21-07-2022).
- ❖ **Hejda, M.**, Malysheva, E., Owen-Newns, D., Al-Taai, Q.R.A., Zhang, W., Ortega-Piwonka, I., Javaloyes, J., Wasige, E., Dolores-Calzadilla, V., Figueiredo, J.M.L., Romeira, B. & Hurtado, A. *Artificial optoelectronic spiking neurons with*

laser-coupled resonant tunnelling diode systems. IEEE International Photonics Conference (IPC), Vancouver, Canada. (11-11-2022).

- ❖ **Hejda, M.**, Doris, E. A., Bilodeau, S., Robertson, J., Xu, L., Shastri, B. J., Prucnal, P. R & Hurtado, A. *Synaptic weighting of spiking VCSEL-neurons using integrated photonic microring weight banks.* CLEO/EQEC Europe, München, Germany. (30-06-2023).

List of references

1. Burgener, E. & Rydning, J. *High Data Growth and Modern Applications Drive New Storage Requirements in Digitally Transformed Enterprises* (IDC, 2022).
2. Kaynak, O. The golden age of Artificial Intelligence. *Discover Artificial Intelligence* **1**, 1. doi:10.1007/s44163-021-00009-x (2021).
3. Leijnen, S. & Veen, F. v. The Neural Network Zoo. *Proceedings* **47**, 9. doi:10.3390/proceedings2020047009 (2020).
4. Maass, W. Networks of spiking neurons: The third generation of neural network models. *Neural Networks* **10**, 1659–1671. doi:10.1016/S0893-6080(97)00011-7 (1997).
5. Minsky, M. & Papert, S. *Perceptrons: An Introduction to Computational Geometry* 292 pp. (MIT Press, Cambridge, MA, USA, 1988).
6. Haber, E. & Ruthotto, L. Stable architectures for deep neural networks. *Inverse Problems* **34**, 014004. doi:10.1088/1361-6420/aa9a90 (2018).
7. Linnainmaa, S. Taylor expansion of the accumulated rounding error. *BIT* **16**, 146–160. doi:10.1007/BF01931367 (1976).
8. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536. doi:10.1038/323533a0 (1986).
9. Huang, C., La Barbera, S. & Small, C. Building brain-inspired computing. *Nature Communications* **10**, 4838. doi:10.1038/s41467-019-12521-x (2019).
10. Bouvier, M. *et al.* Spiking Neural Networks Hardware Implementations and Challenges. *ACM Journal on Emerging Technologies in Computing Systems* **15**, 1–35. doi:10.1145/3304103 (2019).
11. Pfeiffer, M. & Pfeil, T. Deep Learning With Spiking Neurons: Opportunities and Challenges. *Frontiers in Neuroscience* **12**. doi:10.3389/fnins.2018.00774 (October 2018).
12. Sarpeshkar, R. Analog Versus Digital: Extrapolating from Electronics to Neurobiology. *Neural Computation* **10**, 1601–1638. doi:10.1162/089976698300017052 (1998).
13. Lee, C. *et al.* Enabling Spike-Based Backpropagation for Training Deep Neural Network Architectures. *Frontiers in Neuroscience* **14**, 119. doi:10.3389/fnins.2020.00119 (2020).

14. McKennoch, S., Liu, D. & Bushnell, L. G. *Fast Modifications of the SpikeProp Algorithm* in *The 2006 IEEE International Joint Conference on Neural Network Proceedings* (IEEE, 2006), 3970–3977. doi:10.1109/IJCNN.2006.246918.
15. Shrestha, S. B. & Orchard, G. *SLAYER: Spike Layer Error Reassignment in Time* 2018. arXiv: 1810.08646.
16. Ponulak, F. & Kasiński, A. Supervised Learning in Spiking Neural Networks with ReSuMe: Sequence Learning, Classification, and Spike Shifting. *Neural Computation* **22**, 467–510. doi:10.1162/neco.2009.11-08-901 (2010).
17. Zenke, F. & Ganguli, S. SuperSpike: Supervised Learning in Multilayer Spiking Neural Networks. *Neural Computation* **30**, 1514–1541. doi:10.1162/neco_a_01086 (2018).
18. Shrestha, S. B. & Song, Q. Robust spike-train learning in spike-event based weight update. *Neural Networks* **96**, 33–46. doi:10.1016/j.neunet.2017.08.010 (2017).
19. Rathi, N., Srinivasan, G., Panda, P. & Roy, K. *Enabling Deep Spiking Neural Networks with Hybrid Conversion and Spike Timing Dependent Backpropagation* 2020. arXiv: 2005.01807[cs, stat].
20. Bengio, Y. *et al.* STDP-Compatible Approximation of Backpropagation in an Energy-Based Model. *Neural Computation* **29**, 555–577. doi:10.1162/NECO_a_00934 (2017).
21. Martin, E. *et al.* EqSpike: Spike-driven equilibrium propagation for neuromorphic implementations. *iScience* **24**, 102222. doi:10.1016/j.isci.2021.102222 (2021).
22. Strogatz, S. H. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* 1 p. (CRC Press, Boca Raton, FL, 2018).
23. Delvenne, J.-C. What is a universal computing machine? *Applied Mathematics and Computation* **215**, 1368–1374. doi:10.1016/j.amc.2009.04.057 (2009).
24. Appeltant, L. *et al.* Information processing using a single dynamical node as complex system. *Nature Communications* **2**, 466–468. doi:10.1038/ncomms1476 (2011).
25. Dambre, J., Verstraeten, D., Schrauwen, B. & Massar, S. Information Processing Capacity of Dynamical Systems. *Scientific Reports* **2**, 514. doi:10.1038/srep00514 (2012).
26. Nakajima, K. *et al.* A soft body as a reservoir: case studies in a dynamic model of octopus-inspired soft robotic arm. *Frontiers in Computational Neuroscience* **7**. doi:10.3389/fncom.2013.00091 (2013).
27. Bueno, J., Robertson, J., Hejda, M. & Hurtado, A. Comprehensive Performance Analysis of a VCSEL-Based Photonic Reservoir Computer. *IEEE Photonics Technology Letters* **33**, 920–923. doi:10.1109/LPT.2021.3075095 (2021).

28. Guevara, M. R. in *Nonlinear Dynamics in Physiology and Medicine* (eds Beuter, A., Glass, L., Mackey, M. C. & Titcombe, M. S.) 41–85 (Springer, New York, NY, 2003). doi:10.1007/978-0-387-21640-9_3.
29. Zhao, Z. & Gu, H. Transitions between classes of neuronal excitability and bifurcations induced by autapse. *Scientific Reports* **7**, 6760. doi:10.1038/s41598-017-07051-9 (2017).
30. Sneyd, J., Girard, S. & Clapham, D. Calcium wave propagation by calcium-induced calcium release: An unusual excitable system. *Bulletin of Mathematical Biology* **55**, 315–344. doi:10.1016/S0092-8240(05)80268-X (1993).
31. D’Huys, O. *et al.* Canard resonance: on noise-induced ordering of trajectories in heterogeneous networks of slow-fast systems. *Journal of Physics: Photonics* **3**, 024010. doi:10.1088/2515-7647/abcbe3 (2021).
32. Krauskopf, B. *et al.* Excitability and self-pulsations near homoclinic bifurcations in semiconductor laser systems. *Optics Communications* **215**, 367–379. doi:10.1016/S0030-4018(02)02239-3 (2003).
33. Wünsche, H. J., Brox, O., Radziunas, M. & Henneberger, F. Excitability of a Semiconductor Laser by a Two-Mode Homoclinic Bifurcation. *Physical Review Letters* **88**, 4. doi:10.1103/PhysRevLett.88.023901 (2002).
34. Izhikevich, E. M. *Dynamical systems in neuroscience: the geometry of excitability and bursting* 1st Ed. 441 pp. (MIT Press, London, 2007).
35. Hodgkin, A. L. & Huxley, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology* **117**, 500–544. doi:10.1113/jphysiol.1952.sp004764 (1952).
36. Hodgkin, A. L., Huxley, A. F. & Katz, B. Measurement of current-voltage relations in the membrane of the giant axon of *Loligo*. *The Journal of Physiology* **116**, 424–448. doi:10.1113/jphysiol.1952.sp004716 (1952).
37. Morris, C. & Lecar, H. Voltage oscillations in the barnacle giant muscle fiber. *Biophysical Journal* **35**, 193–213. doi:10.1016/S0006-3495(81)84782-0 (1981).
38. Zou, Q. *et al.* Real time simulations of networks of Hodgkin-Huxley neurons using analog circuits in *Annual Computational Neuroscience Meeting, (CNS 2005)* (United States, 2005), 00.
39. Nagumo, J.-i., Arimoto, S. & Yoshizawa, S. An Active Pulse Transmission Line Simulating Nerve Axon. *Proceedings of the IRE* **50**, 2061–2070. doi:10.1109/JRPROC.1962.288235 (1962).
40. Izhikevich, E. M. Resonate-and-fire neurons. *Neural Networks* **14**, 883–894. doi:10.1016/S0893-6080(01)00078-8 (2001).

41. Desroches, M., Krupa, M. & Rodrigues, S. Inflection, canards and excitability threshold in neuronal models. *Journal of Mathematical Biology* **67**, 989–1017. doi:10.1007/s00285-012-0576-z (2013).
42. Izhikevich, E. M. Class I neural excitability, conventional synapses, weakly connected networks, and mathematical foundations of pulse-coupled models. *IEEE Transactions on Neural Networks* **10**, 499–507. doi:10.1109/72.761707 (1999).
43. Dolcemascolo, A. *et al.* Resonator neuron and triggering multipulse excitability in laser with injected signal. *Physical Review E* **98**, 062211. doi:10.1103/PhysRevE.98.062211 (2018).
44. Prescott, S. A. in *Encyclopedia of Computational Neuroscience* (eds Jaeger, D. & Jung, R.) 1–7 (Springer New York, New York, NY, 2014). doi:10.1007/978-1-4614-7320-6_151-1.
45. *GoPhotonics: Laser Diodes* GoPhotonics: Laser Diodes. <https://www.gophotonics.com/search/laser-diodes/filters?page=1&country=global&stechnology=;Fabry-Perot%20%28FP%29;..>
46. Numai, T. *Fundamentals of Semiconductor Lasers* doi:10.1007/978-4-431-55148-5 (Springer Japan, Tokyo, 2015).
47. Liu, Z. & Slavik, R. Optical Injection Locking: From Principle to Applications. *Journal of Lightwave Technology* **38**, 43–59. doi:10.1109/JLT.2019.2945718 (2020).
48. Hwang, S. K. & Liang, D. H. Effects of linewidth enhancement factor on period-one oscillations of optically injected semiconductor lasers. *Applied Physics Letters* **89**, 061120. doi:10.1063/1.2335978 (2006).
49. Weiser, M. D. The Computer for the 21st Century. *Scientific American* **265**, 94–105. doi:10.1038/scientificamerican0991-94 (1991).
50. Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* 2018. arXiv: 1810.04805.
51. Chen, C. *et al.* A Performance Evaluation of Machine Learning-Based Streaming Spam Tweets Detection. *IEEE Transactions on Computational Social Systems* **2**, 65–76. doi:10.1109/TCSS.2016.2516039 (2015).
52. Vinyals, O. *et al.* Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature* **575**, 350–354. doi:10.1038/s41586-019-1724-z (2019).
53. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589. doi:10.1038/s41586-021-03819-2 (2021).
54. Jumper, J. & Hassabis, D. Protein structure predictions to atomic accuracy with AlphaFold. *Nature Methods* **19**, 11–12. doi:10.1038/s41592-021-01362-6 (2022).
55. Ramesh, A. *et al.* *Hierarchical Text-Conditional Image Generation with CLIP Latents* 2022. doi:10.48550/arXiv.2204.06125. arXiv: 2204.06125[cs].

56. Saharia, C. *et al.* *Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding* 2022. doi:10.48550/arXiv.2205.11487. arXiv: 2205.11487[cs].
57. Rombach, R. *et al.* *High-Resolution Image Synthesis with Latent Diffusion Models* 2022. doi:10.48550/arXiv.2112.10752. arXiv: 2112.10752[cs].
58. Brown, T. B. *et al.* *Language Models are Few-Shot Learners* 2020. arXiv: 2005.14165.
59. Castelvechi, D. Are ChatGPT and AlphaCode going to replace programmers? *Nature*. doi:10.1038/d41586-022-04383-z (2022).
60. Haque, M. U. *et al.* "I think this is the most disruptive technology": Exploring Sentiments of ChatGPT Early Adopters using Twitter Data 2022. doi:10.48550/arXiv.2212.05856. arXiv: 2212.05856[cs].
61. Katz, D. M., Bommarito, M. J., Gao, S. & Arredondo, P. GPT-4 Passes the Bar Exam. *SSRN Electronic Journal*. doi:10.2139/ssrn.4389233 (2023).
62. Apple. *Hey Siri: An On-device DNN-powered Voice Trigger for Apple's Personal Assistant* Apple Machine Learning Blog. <https://machinelearning.apple.com/research/hey-siri> (01/05/2020).
63. Lane, N. D. *et al.* Squeezing Deep Learning into Mobile and Embedded Devices. *IEEE Pervasive Computing* **16**, 82–88. doi:10.1109/MPRV.2017.2940968 (2017).
64. LeCun, Y. *Deep Learning Hardware: Past, Present, and Future in 2019 IEEE International Solid-State Circuits Conference - (ISSCC)* (IEEE, 2019), 12–19. doi:10.1109/ISSCC.2019.8662396.
65. Lohn, A. J. & Musser, M. *AI & Compute: How Much Longer Can Computing Power Drive Artificial Intelligence Progress?* (CSET (Centre for Security and Emerging Technology), Georgetown University, 2022).
66. Li, C. *OpenAI's GPT-3 Language Model: A Technical Overview* 2020.
67. Anthony, L. F. W., Kanding, B. & Selvan, R. *Carbontracker: Tracking and Predicting the Carbon Footprint of Training Deep Learning Models* 2020. doi:10.48550/arXiv.2007.03051. arXiv: 2007.03051[cs, eess, stat].
68. Strubell, E., Ganesh, A. & McCallum, A. *Energy and Policy Considerations for Deep Learning in NLP* 2019. arXiv: 1906.02243.
69. So, D. R., Liang, C. & Le, Q. V. *The Evolved Transformer* 2019. doi:10.48550/arXiv.1901.11117. arXiv: 1901.11117[cs, stat].
70. Quach, K. *AI me to the Moon... Carbon footprint for 'training GPT-3' same as driving to our natural satellite and back* https://www.theregister.com/2020/11/04/gpt3_carbon_footprint_estimate/ (01/09/2023).

71. Attwell, D. & Laughlin, S. B. An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism* **21**, 1133–1145. doi:10.1097/00004647-200110000-00001 (2001).
72. Kozachkov, L., Kastanenko, K. V. & Krotov, D. *Building Transformers from Neurons and Astrocytes* 2022. doi:10.1101/2022.10.12.511910.
73. Dennard, R. H. *et al.* Design of ion-implanted MOSFET's with very small physical dimensions. *IEEE Journal of Solid-State Circuits* **9**, 256–268. doi:10.1109/JSSC.1974.1050511 (1974).
74. Koomey, J., Berard, S., Sanchez, M. & Wong, H. Implications of Historical Trends in the Electrical Efficiency of Computing. *IEEE Annals of the History of Computing* **33**, 46–54. doi:10.1109/MAHC.2010.28 (2011).
75. Peper, F. The End of Moore's Law: Opportunities for Natural Computing? *New Generation Computing* **35**, 253–269. doi:10.1007/s00354-017-0020-4 (2017).
76. Mack, C. The Multiple Lives of Moore's Law. *IEEE Spectrum* **52**, 31–31. doi:10.1109/MSPEC.2015.7065415 (2015).
77. Rupp, K. *Microprocessor Trend Data (source code)* 2023. <https://github.com/karlrupp/microprocessor-trend-data> (03/19/2023).
78. Wang, Z. *et al.* Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. *Nature Materials* **16**, 101–108. doi:10.1038/nmat4756 (2017).
79. Boybat, I. *et al.* Neuromorphic computing with multi-memristive synapses. *Nature Communications* **9**, 1–12. doi:10.1038/s41467-018-04933-y (2018).
80. Covi, E. *et al.* Ferroelectric-based synapses and neurons for neuromorphic computing. *Neuromorphic Computing and Engineering* **2**, 012002. doi:10.1088/2634-4386/ac4918 (2022).
81. Zahedinejad, M. *et al.* Two-dimensional mutually synchronized spin Hall nano-oscillator arrays for neuromorphic computing. *Nature Nanotechnology* **15**, 47–52. doi:10.1038/s41565-019-0593-9 (2020).
82. Farkhani, H. *et al.* LAO-NCS: Laser Assisted Spin Torque Nano Oscillator-Based Neuromorphic Computing System. *Frontiers in Neuroscience* **13**. doi:10.3389/fnins.2019.01429 (2020).
83. Hylton, T. Thermodynamic Neural Network. *Entropy* **22**, 256. doi:10.3390/e22030256 (2020).
84. Hylton, T. Thermodynamic State Machine Network. *Entropy* **24**, 744. doi:10.3390/e24060744 (2022).
85. Rudé, M. *et al.* Interferometric photodetection in silicon photonics for phase diffusion quantum entropy sources. *Optics Express* **26**, 31957. doi:10.1364/OE.26.031957 (2018).

86. Brito, J. P. *et al.* *Quantum Services Architecture in Softwarized Infrastructures in 2019 21st International Conference on Transparent Optical Networks (ICTON) 2019 21st International Conference on Transparent Optical Networks (ICTON)* (IEEE, Angers, France, 2019), 1–4. doi:10.1109/ICTON.2019.8840400.
87. Mead, C. *Analog VLSI and neural systems* (Addison-Wesley, Reading, MA., 1989).
88. Grand View Research. *Neuromorphic Computing Market Size, Share & Trends Analysis Report By Application, By End-Use, By Deployment, By Component, By Region, And Segment Forecasts, 2023 - 2030* GVR-1-68038-066-8 (2022).
89. Mordor Intelligence. *Neuromorphic Chip Market - Growth, Trends, COVID-19 Impact, and Forecasts (2023 - 2028)* Market Forecast (India, 2022).
90. Bose, S. K., Acharya, J. & Basu, A. *Is my Neural Network Neuromorphic? Taxonomy, Recent Trends and Future Directions in Neuromorphic Engineering* 2020. arXiv: 2002.11945.
91. Indiveri, G. *et al.* Neuromorphic Silicon Neuron Circuits. *Frontiers in Neuroscience* **5**. doi:10.3389/fnins.2011.00073 (2011).
92. Pisarev, A., Busygin, A., Udovichenko, S. Y. & Maevsky, O. A biomorphic neuroprocessor based on a composite memristor-diode crossbar. *Microelectronics Journal* **102**, 104827. doi:10.1016/j.mejo.2020.104827 (2020).
93. Davies, M. *et al.* Loihi: A Neuromorphic Manycore Processor with On-Chip Learning. *IEEE Micro* **38**, 82–99. doi:10.1109/MM.2018.112130359 (2018).
94. Schuman, C. D. *et al.* *A Survey of Neuromorphic Computing and Neural Networks in Hardware* 2017. arXiv: 1705.06963.
95. Rhodes, O. *et al.* Real-time cortical simulation on neuromorphic hardware. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **378**, 20190160. doi:10.1098/rsta.2019.0160 (2020).
96. Gaba, S. *et al.* Stochastic memristive devices for computing and neuromorphic applications. *Nanoscale* **5**, 5872. doi:10.1039/c3nr01176c (2013).
97. Gallego, G. *et al.* *Event-based Vision: A Survey* 2019. arXiv: 1904.08405.
98. Lungu, I. A., Liu, S.-C. & Delbruck, T. Incremental Learning of Hand Symbols Using Event-Based Cameras. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* **9**, 690–696. doi:10.1109/JETCAS.2019.2951062 (2019).
99. Lagorce, X. *et al.* Asynchronous Event-Based Multikernel Algorithm for High-Speed Visual Features Tracking. *IEEE Transactions on Neural Networks and Learning Systems* **26**, 1710–1720. doi:10.1109/TNNLS.2014.2352401 (2015).

100. Gehrig, M., Shrestha, S. B., Mouritzen, D. & Scaramuzza, D. *Event-Based Angular Velocity Regression with Spiking Networks* in *2020 IEEE International Conference on Robotics and Automation (ICRA)* 2020 IEEE International Conference on Robotics and Automation (ICRA) (IEEE, Paris, France, 2020), 4195–4202. doi:10.1109/ICRA40945.2020.9197133.
101. Benjamin, B. V. *et al.* Neurogrid: A Mixed-Analog-Digital Multichip System for Large-Scale Neural Simulations. *Proceedings of the IEEE* **102**, 699–716. doi:10.1109/JPROC.2014.2313565 (2014).
102. Höppner, S. *et al.* *The SpiNNaker 2 Processing Element Architecture for Hybrid Digital Neuromorphic Computing* 2021. arXiv: 2103.08392.
103. Mayr, C., Hoepfner, S. & Furber, S. *SpiNNaker 2: A 10 Million Core Processor System for Brain Simulation and Machine Learning* 2019. arXiv: 1911.02385.
104. Yan, Y. *et al.* Comparing Loihi with a SpiNNaker 2 prototype on low-latency keyword spotting and adaptive robotic control. *Neuromorphic Computing and Engineering* **1**, 014002. doi:10.1088/2634-4386/abf150 (2021).
105. Diamond, A., Nowotny, T. & Schmuker, M. Comparing Neuromorphic Solutions in Action: Implementing a Bio-Inspired Solution to a Benchmark Classification Task on Three Parallel-Computing Platforms. *Frontiers in Neuroscience* **9**. doi:10.3389/fnins.2015.00491 (2016).
106. Cramer, B. *et al.* Surrogate gradients for analog neuromorphic computing. *Proceedings of the National Academy of Sciences* **119**, e2109194119. doi:10.1073/pnas.2109194119 (2022).
107. Frenkel, C. & Indiveri, G. *ReckOn: A 28nm Sub-mm² Task-Agnostic Spiking Recurrent Neural Network Processor Enabling On-Chip Learning over Second-Long Timescales* in *2022 IEEE International Solid-State Circuits Conference (ISSCC)* (2022), 1–3. doi:10.1109/ISSCC42614.2022.9731734. arXiv: 2208.09759[cs].
108. Pei, J. *et al.* Towards artificial general intelligence with hybrid Tianjic chip architecture. *Nature* **572**, 106–111. doi:10.1038/s41586-019-1424-8 (2019).
109. Davies, M. *et al.* Advancing Neuromorphic Computing With Loihi: A Survey of Results and Outlook. *Proceedings of the IEEE* **109**, 911–934. doi:10.1109/JPROC.2021.3067593 (2021).
110. Orchard, G. *et al.* *Efficient Neuromorphic Signal Processing with Loihi 2* in *2021 IEEE Workshop on Signal Processing Systems (SiPS)* 2021 IEEE Workshop on Signal Processing Systems (SiPS) (IEEE, Coimbra, Portugal, 2021), 254–259. doi:10.1109/SiPS52927.2021.00053.
111. DeBole, M. V. *et al.* TrueNorth: Accelerating From Zero to 64 Million Neurons in 10 Years. *Computer* **52**, 20–29. doi:10.1109/MC.2019.2903009 (2019).

112. Cassidy, A. S. *et al.* *Cognitive computing building block: A versatile and efficient digital neuron model for neurosynaptic cores* in *The 2013 International Joint Conference on Neural Networks (IJCNN)* The 2013 International Joint Conference on Neural Networks (IJCNN) (2013), 1–10. doi:10.1109/IJCNN.2013.6707077.
113. Vanarse, A., Osseiran, A., Rassau, A. & van der Made, P. A Hardware-Deployable Neuromorphic Solution for Encoding and Classification of Electronic Nose Data. *Sensors* **19**, 4831. doi:10.3390/s19224831 (2019).
114. Zins, N., Zhang, Y., Yu, C. & An, H. in *Frontiers of Quality Electronic Design (QED): AI, IoT and Hardware Security* (ed Iranmanesh, A.) 259–296 (Springer International Publishing, Cham, 2023). doi:10.1007/978-3-031-16344-9_7.
115. Houshmand, P. *et al.* DIANA: An End-to-End Hybrid Digital and Analog Neural Network SoC for the Edge. *IEEE Journal of Solid-State Circuits* **58**, 203–215. doi:10.1109/JSSC.2022.3214064 (2023).
116. Wan, W. *et al.* A compute-in-memory chip based on resistive random-access memory. *Nature* **608**, 504–512. doi:10.1038/s41586-022-04992-8 (2022).
117. Furber, S. B. Large-scale neuromorphic computing systems. *Journal of Neural Engineering* **13**, 051001. doi:10.1088/1741-2560/13/5/051001 (2016).
118. Ivanov, D. *et al.* *Neuromorphic Artificial Intelligence Systems 2022*. arXiv: 2205.13037[cs].
119. Liu, Q. *et al.* *Live Demonstration: Face Recognition on an Ultra-Low Power Event-Driven Convolutional Neural Network ASIC* in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (IEEE, Long Beach, CA, USA, 2019), 1680–1681. doi:10.1109/CVPRW.2019.00213.
120. Bekolay, T. *et al.* Nengo: a Python tool for building large-scale functional brain models. *Frontiers in Neuroinformatics* **7**. doi:10.3389/fninf.2013.00048 (2014).
121. Mundy, A., Knight, J., Stewart, T. C. & Furber, S. *An efficient SpiNNaker implementation of the Neural Engineering Framework* in *2015 International Joint Conference on Neural Networks (IJCNN)* 2015 International Joint Conference on Neural Networks (IJCNN) (2015), 1–8. doi:10.1109/IJCNN.2015.7280390.
122. Mack, J. *et al.* RANC: Reconfigurable Architecture for Neuromorphic Computing. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* **40**, 2265–2278. doi:10.1109/TCAD.2020.3038151 (2021).
123. Hazan, H. *et al.* BindsNET: A Machine Learning-Oriented Spiking Neural Networks Library in Python. *Frontiers in Neuroinformatics* **12**. doi:10.3389/fninf.2018.00089 (2018).

124. Eshraghian, J. K. *et al.* Training Spiking Neural Networks Using Lessons From Deep Learning. *Proceedings of the IEEE* **111**, 1016–1054. doi:10.1109/JPROC.2023.3308088 (2023).
125. Pehle, C.-G. & Pedersen, J. E. *Norse - A deep learning library for spiking neural networks* version 0.0.5. 2021. doi:10.5281/ZENODO.4422025. <https://zenodo.org/record/4422025> (01/22/2023).
126. Fang, W. *et al.* *SpikingJelly: Open-source Deep Learning SNN Framework* <https://github.com/fangwei123456/spikingjelly>.
127. Lenz, G. *et al.* *Tonic: event-based datasets and transformations*. version 0.4.0. 2021. doi:10.5281/zenodo.5079802. <https://doi.org/10.5281/zenodo.5079802>.
128. UZH Robotics and Perception Group. *University of Zurich: Event-based Vision Resources repository* https://github.com/uzh-rpg/event-based_vision_resources#learning-methods-frameworks (01/22/2023).
129. Psaltis, D. & Farhat, N. Optical information processing based on an associative-memory model of neural nets with thresholding and feedback. *Optics Letters* **10**, 98–100. doi:10.1364/OL.10.000098 (1985).
130. Kiamilev, F. *et al.* Programmable Optoelectronic Multiprocessors And Their Comparison With Symbolic Substitution For Digital Optical Computing. *Optical Engineering* **28**. doi:10.1117/12.7976967 (1989).
131. Waterson, C. & Jenkins, B. K. Shared-memory optical/electronic computer: architecture and control. *Applied Optics* **33**, 1559. doi:10.1364/AO.33.001559 (1994).
132. Moerland, P. D., Fiesler, E. & Saxena, I. Incorporation of liquid-crystal light valve nonlinearities in optical multilayer neural networks. *Applied Optics* **35**, 5301. doi:10.1364/AO.35.005301 (1996).
133. Mos, E. C. *et al.* Optical neuron by use of a laser diode with injection seeding and external optical feedback. *IEEE Transactions on Neural Networks* **11**, 988–996. doi:10.1109/72.857778 (2000).
134. Rosenbluth, D., Kravtsov, K., Fok, M. P. & Prucnal, P. R. A high performance photonic pulse processing device. *Optics Express* **17**, 22767. doi:10.1364/OE.17.022767 (2009).
135. Fok, M. P. *et al.* Signal feature recognition based on lightwave neuromorphic signal processing. *Optics Letters* **36**, 19. doi:10.1364/OL.36.000019 (2011).
136. Kravtsov, K. S., Fok, M. P., Prucnal, P. R. & Rosenbluth, D. Ultrafast All-Optical Implementation of a Leaky Integrate-and-Fire Neuron. *Optics Express* **19**, 2133. doi:10.1364/OE.19.002133 (2011).
137. Margalit, N. *et al.* Perspective on the future of silicon photonics and electronics. *Applied Physics Letters* **118**, 220501. doi:10.1063/5.0050117 (2021).
138. Shastri, B. J. *et al.* Photonics for artificial intelligence and neuromorphic computing. *Nature Photonics* **15**, 102–114. doi:10.1038/s41566-020-00754-y (2021).

139. Ren, H., Li, X., Zhang, Q. & Gu, M. On-chip noninterference angular momentum multiplexing of broadband light. *Science* **352**, 805–809. doi:10.1126/science.aaf1112 (2016).
140. Ferreira de Lima, T. *et al.* Progress in neuromorphic photonics. *Nanophotonics* **6**, 577–599. doi:10.1515/nanoph-2016-0139 (2017).
141. Shen, Y. *et al.* Deep learning with coherent nanophotonic circuits. *Nature Photonics* **11**, 441. doi:10.1038/nphoton.2017.93 (2017).
142. Miller, D. A. B. Attojoule Optoelectronics for Low-Energy Information Processing and Communications. *Journal of Lightwave Technology* **35**, 346–396. doi:10.1109/JLT.2017.2647779 (2017).
143. Jutamulia, S. & Yu, F. T. S. Overview of hybrid optical neural networks. *Optics & Laser Technology* **28**, 59–72. doi:10.1016/0030-3992(95)00070-4 (1996).
144. Caulfield, J. H. & Dolev, S. Why future supercomputing requires optics. *Nature Photonics* **4**, 261–263. doi:10.1038/nphoton.2010.94 (2010).
145. Zhou, L., Shi, J. & Zhang, X. The mr-MDA: An Invariant to Shifting, Scaling, and Rotating Variance for 3D Object Recognition Using Diffractive Deep Neural Network. *Sensors* **22**, 7754. doi:10.3390/s22207754 (2022).
146. Shastri, B. J. *et al.* *Neuromorphic Photonic Processor Applications (Invited Talk) in 2019 IEEE Photonics Society Summer Topical Meeting Series (SUM)* (IEEE, 2019), 1–2. doi:10.1109/PHOSST.2019.8795013.
147. Ríos Ocampo, C. A. *et al.* In-memory computing on a photonic platform. *Science Advances* **5**, eaau5759. doi:10.1126/sciadv.aau5759 (2019).
148. Liu, W. *et al.* *HolyLight: A Nanophotonic Accelerator for Deep Learning in Data Centers* in *2019 Design, Automation Test in Europe Conference Exhibition (DATE) 2019 Design, Automation Test in Europe Conference Exhibition (DATE)* (2019), 1483–1488. doi:10.23919/DATE.2019.8715195.
149. Marković, D., Mizrahi, A., Querlioz, D. & Grollier, J. Physics for neuromorphic computing. *Nature Reviews Physics* **2**, 499–510. doi:10.1038/s42254-020-0208-2 (2020).
150. Romeira, B., Figueiredo, J. M. L. & Javaloyes, J. NanoLEDs for energy-efficient and gigahertz-speed spike-based sub- neuromorphic nanophotonic computing. *Nanophotonics*. doi:10.1515/nanoph-2020-0177 (2020).
151. Pai, S. *et al.* *Parallel fault-tolerant programming of an arbitrary feedforward photonic network* 2019. arXiv: 1909.06179.
152. Lee, Y.-J. *et al.* Photonic spiking neural networks with event-driven femtojoule optoelectronic neurons based on Izhikevich-inspired model. *Optics Express* **30**, 19360. doi:10.1364/OE.449528 (2022).

153. Shainline, J. M. *The Largest Cognitive Systems Will be Optoelectronic in 2018 IEEE International Conference on Rebooting Computing (ICRC)* (IEEE, 2018), 1–10. doi:10.1109/ICRC.2018.8638599.
154. De Marinis, L., Cococcioni, M., Castoldi, P. & Andriolli, N. Photonic Neural Networks: A Survey. *IEEE Access* **7**, 175827–175841. doi:10.1109/ACCESS.2019.2957245 (2019).
155. Stabile, R. *et al.* Neuromorphic photonics: 2D or not 2D? *Journal of Applied Physics* **129**, 200901. doi:10.1063/5.0047946 (2021).
156. Du, Y. *et al.* Implementation of optical neural network based on Mach–Zehnder interferometer array. *IET Optoelectronics* **17**. doi:10.1049/ote2.12086 (2022).
157. Reck, M., Zeilinger, A., Bernstein, H. J. & Bertani, P. Experimental realization of any discrete unitary operator. *Physical Review Letters* **73**, 58–61. doi:10.1103/PhysRevLett.73.58 (1994).
158. Shokraneh, F., Geoffroy-Gagnon, S., Nezami, M. S. & Liboiron-Ladouceur, O. A Single Layer Neural Network Implemented by a 4x4 MZI-Based Optical Processor. *IEEE Photonics Journal* **11**, 1–12. doi:10.1109/JPHOT.2019.2952562 (2019).
159. Hughes, T. W., Minkov, M., Shi, Y. & Fan, S. Training of photonic neural networks through in situ backpropagation and gradient measurement. *Optica* **5**, 864–871. doi:10.1364/OPTICA.5.000864 (2018).
160. Pai, S. *et al.* *Experimentally realized in situ backpropagation for deep learning in nanophotonic neural networks* 2022. arXiv: 2205.08501[physics].
161. Shao, R., Zhang, G. & Gong, X. Generalized robust training scheme using genetic algorithm for optical neural networks with imprecise components. *Photonics Research* **10**, 1868. doi:10.1364/PRJ.449570 (2022).
162. Williamson, I. A. D. *et al.* Reprogrammable Electro-Optic Nonlinear Activation Functions for Optical Neural Networks. *IEEE Journal of Selected Topics in Quantum Electronics* **26**, 1–12. doi:10.1109/JSTQE.2019.2930455 (2020).
163. Fard, M. M. P. *et al.* Experimental realization of arbitrary activation functions for optical neural networks. *Optics Express* **28**, 12138–12148. doi:10.1364/OE.391473 (2020).
164. Tian, Y. *et al.* Scalable and compact photonic neural chip with low learning-capability-loss. *Nanophotonics* **11**, 329–344. doi:10.1515/nanoph-2021-0521 (2022).
165. Shafiee, A. *et al.* *LoCI: An Analysis of the Impact of Optical Loss and Crosstalk Noise in Integrated Silicon-Photonic Neural Networks* 2022. arXiv: 2204.03835.
166. Xu, X. *et al.* Self-calibrating programmable photonic integrated circuits. *Nature Photonics* **16**, 595–602. doi:10.1038/s41566-022-01020-z (2022).
167. Xu, X. *et al.* Photonic Perceptron Based on a Kerr Microcomb for High-Speed, Scalable, Optical Neural Networks. *Laser & Photonics Reviews* **14**, 2000070. doi:10.1002/lpor.202000070 (2020).

168. Mourgias-Alexandris, G. *et al.* An all-optical neuron with sigmoid activation function. *Optics Express* **27**, 9620. doi:10.1364/OE.27.009620 (2019).
169. Shi, B., Calabretta, N. & Stabile, R. Deep Neural Network Through an InP SOA-Based Photonic Integrated Cross-Connect. *IEEE Journal of Selected Topics in Quantum Electronics* **26**, 1–11. doi:10.1109/JSTQE.2019.2945548 (2020).
170. Toole, R. *et al.* Photonic Implementation of Spike-Timing-Dependent Plasticity and Learning Algorithms of Biological Neural Systems. *Journal of Lightwave Technology* **34**, 470–476. doi:10.1109/JLT.2015.2475275 (2016).
171. Yang, L., Zhang, L. & Ji, R. *On-chip optical matrix-vector multiplier* in. SPIE Optical Engineering + Applications (eds Iftekharuddin, K. M., Awwal, A. A. S. & Márquez, A.) (San Diego, California, United States, 2013), 88550F. doi:10.1117/12.2028585.
172. Tait, A. N. *et al.* Silicon Photonic Modulator Neuron. *Physical Review Applied* **11**, 064043. doi:10.1103/PhysRevApplied.11.064043 (2019).
173. Tait, A. N., Nahmias, M. A., Shastri, B. J. & Prucnal, P. R. Broadcast and Weight: An Integrated Network For Scalable Photonic Spike Processing. *Journal of Lightwave Technology* **32**, 4029–4041. doi:10.1109/JLT.2014.2345652 (2014).
174. Huang, C. *et al.* Demonstration of scalable microring weight bank control for large-scale photonic integrated circuits. *APL Photonics* **5**, 040803. doi:10.1063/1.5144121 (2020).
175. Xu, S., Wang, J. & Zou, W. *Performance evaluation of an integrated photonic convolutional neural network based on delay buffering and wavelength division multiplexing* 2020. arXiv: 1910.12635v2.
176. Xu, L. *et al.* Scalable Networks of Neuromorphic Photonic Integrated Circuits. *IEEE Journal of Selected Topics in Quantum Electronics* **28**, 1–9. doi:10.1109/JSTQE.2022.3211453 (2022).
177. Gelens, L. *et al.* Excitability in semiconductor microring lasers: Experimental and theoretical pulse characterization. *Physical Review A* **82**, 063841. doi:10.1103/PhysRevA.82.063841 (2010).
178. Coomans, W. *et al.* Solitary and coupled semiconductor ring lasers as optical spiking neurons. *Physical Review E* **84**, 036209. doi:10.1103/PhysRevE.84.036209 (2011).
179. Van Vaerenbergh, T., Fiers, M., Dambre, J. & Bienstman, P. Simplified description of self-pulsation and excitability by thermal and free-carrier effects in semiconductor microcavities. *Physical Review A* **86**, 063808. doi:10.1103/PhysRevA.86.063808 (2012).
180. Van Vaerenbergh, T. *et al.* Cascadable excitability in microrings. *Optics Express* **20**, 20292. doi:10.1364/OE.20.020292 (2012).
181. Xiang, J., Torchy, A., Guo, X. & Su, Y. All-Optical Spiking Neuron Based on Passive Microresonator. *Journal of Lightwave Technology* **38**, 4019–4029. doi:10.1109/JLT.2020.2986233 (2020).

182. Yelo-Sarrión, J., Leo, F., Gorza, S.-P. & Parra-Rivas, P. Neuron-like spiking dynamics in asymmetrically driven dissipative nonlinear photonic dimers. *Physical Review A* **106**, 013512. doi:10.1103/PhysRevA.106.013512 (2022).
183. Han, Y. *et al.* An All-MRR-Based Photonic Spiking Neural Network for Spike Sequence Learning. *Photonics* **9**, 120. doi:10.3390/photonics9020120 (2022).
184. George, J. K. *et al.* Neuromorphic photonics with electro-absorption modulators. *Optics Express* **27**, 5181. doi:10.1364/OE.27.005181. arXiv: 1809.03545 (2019).
185. Amin, R. *et al.* An ITO - Graphene heterojunction integrated absorption modulator on Si-photonics for neuromorphic nonlinear activation. *APL Photonics*, 5.0062830. doi:10.1063/5.0062830 (2021).
186. Tsakyridis, A. *et al.* Silicon Photonic Neuromorphic Computing with 16 GHz Input Data and Weight Update Line Rates in. 2022 Conference on Lasers and Electro-Optics (CLEO) (Optica, San Jose, California, 2022).
187. Chakraborty, I., Saha, G., Sengupta, A. & Roy, K. Toward Fast Neural Computing using All-Photonic Phase Change Spiking Neurons. *Scientific Reports* **8**, 12980. doi:10.1038/s41598-018-31365-x (2018).
188. Feldmann, J. *et al.* All-optical spiking neurosynaptic networks with self-learning capabilities. *Nature* **569**, 208–214. doi:10.1038/s41586-019-1157-8 (2019).
189. Feldmann, J. *et al.* Parallel convolutional processing using an integrated photonic tensor core. *Nature* **589**, 52–58. doi:10.1038/s41586-020-03070-1 (2021).
190. Tan, J. Y. S. *et al.* Monadic Pavlovian associative learning in a backpropagation-free photonic network. *Optica* **9**, 792. doi:10.1364/OPTICA.455864 (2022).
191. Burr, G. W. A role for optics in AI hardware. *Nature* **569**, 199–200. doi:10.1038/d41586-019-01406-0 (2019).
192. Lubkin, G. B. Nobel Prize shared by Esaki, Giaever and Josephson. *Physics Today* **26**, 73–75. doi:10.1063/1.3128378 (1973).
193. Shainline, J. M., Buckley, S. M., Mirin, R. P. & Nam, S. W. Superconducting Optoelectronic Circuits for Neuromorphic Computing. *Physical Review Applied* **7**, 034013. doi:10.1103/PhysRevApplied.7.034013 (2017).
194. Schneider, M. *et al.* Supermind: a survey of the potential of superconducting electronics for neuromorphic computing. *Superconductor Science and Technology*. doi:10.1088/1361-6668/ac4cd2 (2022).
195. Zarei, S., Marzban, M.-r. & Khavasi, A. Integrated photonic neural network based on silicon metalines. *Optics Express* **28**, 36668. doi:10.1364/OE.404386 (2020).

196. Wan, L. *et al.* Optical analog computing of spatial differentiation and edge detection with dielectric metasurfaces. *Optics Letters* **45**, 2070. doi:10.1364/OL.386986 (2020).
197. Khoram, E. *et al.* Nanophotonic media for artificial neural inference. *Photonics Research* **7**, 823–827. doi:10.1364/PRJ.7.000823 (2019).
198. Wu, Z. *et al.* Neuromorphic metasurface. *Photonics Research* **8**, 46–50. doi:10.1364/PRJ.8.000046 (2020).
199. Zhou, T. *et al.* Large-scale neuromorphic optoelectronic computing with a reconfigurable diffractive processing unit. *Nature Photonics* **15**, 367–373. doi:10.1038/s41566-021-00796-w (2021).
200. Lin, X. *et al.* All-optical machine learning using diffractive deep neural networks. *Science* **361**, 1004–1008. doi:10.1126/science.aat8084 (2018).
201. Sarantoglou, G., Skontrinis, M. & Mesaritakis, C. All Optical Integrate and Fire Neuromorphic Node Based on Single Section Quantum Dot Laser. *IEEE Journal of Selected Topics in Quantum Electronics* **26**, 1–10. doi:10.1109/JSTQE.2019.2945549 (2020).
202. Robertson, J., Ackemann, T., Lester, L. F. & Hurtado, A. Externally-Triggered Activation and Inhibition of Optical Pulsating Regimes in Quantum-Dot Mode-locked Lasers. *Scientific Reports* **8**, 12515. doi:10.1038/s41598-018-30758-2 (2018).
203. Hayat, A., Bacou, A., Rissons, A. & Mollier, J.-C. in *Advances in Optical and Photonic Devices* (ed Young, K.) (InTech, 2010). doi:10.5772/7141.
204. Michalzik, R. *VCSELs: Fundamentals, Technology and Applications of Vertical-Cavity Surface-Emitting Lasers* 1st ed. 560 pp. (Springer, Heidelberg, 2013).
205. Chang-Hasnain, C. J. *et al.* Dynamic, polarization, and transverse mode characteristics of vertical cavity surface emitting lasers. *IEEE Journal of Quantum Electronics* **27**, 1402–1409. doi:10.1109/3.89957 (1991).
206. Al-Seyab, R. K., Adams, M. J. & Henning, I. D. Dynamics and polarization of conventional and spin-VCSELs in the presence of an axial magnetic field. *Journal of the Optical Society of America B* **32**, 683. doi:10.1364/JOSAB.32.000683 (2015).
207. Ohtsubo, J. *Semiconductor Lasers: Stability, Instability and Chaos* 4th Editio (ed Rhodes, W. T.) 666 pp. doi:10.1007/978-3-319-56138-7 (Springer International Publishing, Cham, 2017).
208. Cheng, H.-T., Yang, Y.-C., Liu, T.-H. & Wu, C.-H. Recent Advances in 850 nm VCSELs for High-Speed Interconnects. *Photonics* **9**, 107. doi:10.3390/photonics9020107 (2022).
209. Chorchos, L. *et al.* Energy Efficient 850 nm VCSEL Based Optical Transmitter and Receiver Link Capable of 80 Gbit/s NRZ Multi-Mode Fiber Data Transmission. *Journal of Lightwave Technology* **38**, 1747–1752. doi:10.1109/JLT.2020.2970299 (2020).

210. GEOLA. *Water Cooled Laser Gain Modules* Water Cooled Laser Gain Modules. <https://www.geola.com/product/water-cooled-dpss-laser-gain-modules/>.
211. Razavi, B. A study of injection locking and pulling in oscillators. *IEEE Journal of Solid-State Circuits* **39**, 1415–1424. doi:10.1109/JSSC.2004.831608 (2004).
212. Chih-Hao Chang, Chrostowski, L. & Chang-Hasnain, C. Injection locking of VCSELs. *IEEE Journal of Selected Topics in Quantum Electronics* **9**, 1386–1393. doi:10.1109/JSTQE.2003.819510 (2003).
213. Murakami, A., Kawashima, K. & Atsuki, K. Cavity resonance shift and bandwidth enhancement in semiconductor lasers with strong light injection. *IEEE Journal of Quantum Electronics* **39**, 1196–1204. doi:10.1109/JQE.2003.817583 (2003).
214. Mogensen, F., Olesen, H. & Jacobsen, G. Locking conditions and stability properties for a semiconductor laser with external light injection. *IEEE Journal of Quantum Electronics* **21**, 784–793. doi:10.1109/JQE.1985.1072760 (1985).
215. Tseng, C.-H. *et al.* High-entropy chaos generation using semiconductor lasers subject to intensity-modulated optical injection for certified physical random number generation. *Optics Letters* **46**, 3384. doi:10.1364/OL.431054 (2021).
216. Simpson, T. B., Liu, J. M., Huang, K. F. & Tai, K. Nonlinear dynamics induced by external optical injection in semiconductor lasers. *Quantum and Semiclassical Optics: Journal of the European Optical Society Part B* **9**, 765–784. doi:10.1088/1355-5111/9/5/009 (1997).
217. Pan, Z. G. *et al.* Optical injection induced polarization bistability in vertical-cavity surface-emitting lasers. *Applied Physics Letters* **63**, 2999–3001. doi:10.1063/1.110264 (1993).
218. Hurtado, A., Henning, I. D. & Adams, M. J. Two-Wavelength Switching With a 1550 nm VCSEL Under Single Orthogonal Optical Injection. *IEEE Journal of Selected Topics in Quantum Electronics* **14**, 911–917. doi:10.1109/JSTQE.2008.918312 (2008).
219. Guo, P. *et al.* Experimental and theoretical study of wide hysteresis cycles in 1550 nm VCSELs under optical injection. *Optics Express* **21**, 3125. doi:10.1364/OE.21.003125 (2013).
220. Hurtado, A. *et al.* Nonlinear dynamics induced by parallel and orthogonal optical injection in 1550 nm Vertical-Cavity Surface-Emitting Lasers (VCSELs). *Optics Express* **18**, 9423–9428. doi:10.1364/OE.18.009423 (2010).
221. Pérez, P., Quirce, A., Pesquera, L. & Valle, A. *Polarization-resolved nonlinear dynamics in long-wavelength single-mode VCSELs subject to orthogonal optical injection* in. SPIE OPTO (eds Witzigmann, B., Henneberger, F., Arakawa, Y. & Freundlich, A.) (San Francisco, California, 2011), 79330B. doi:10.1117/12.875191.

222. Baesens, C. & MacKay, R. S. Interaction of two systems with saddle-node bifurcations on invariant circles: I. Foundations and the mutualistic case. *Nonlinearity* **26**, 3043–3076. doi:10.1088/0951-7715/26/12/3043 (2013).
223. Wieczorek, S., Krauskopf, B., Simpson, T. & Lenstra, D. The dynamical complexity of optically injected semiconductor lasers. *Physics Reports* **416**, 1–128. doi:10.1016/j.physrep.2005.06.003 (2005).
224. Hurtado, A., Henning, I. D. & Adams, M. J. Optical neuron using polarisation switching in a 1550nm-VCSEL. *Optics Express* **18**, 25170–25176. doi:10.1364/OE.18.025170 (2010).
225. Hurtado, A., Schires, K., Henning, I. D. & Adams, M. J. Investigation of vertical cavity surface emitting laser dynamics for neuromorphic photonic systems. *Applied Physics Letters* **100**, 3–7. doi:10.1063/1.3692726 (2012).
226. Al-Seyab, R., Henning, I. D., Adams, M. J. & Hurtado, A. *Controlled Single- and Multiple-Pulse Excitability in VCSELs for Novel Spiking Photonic Neurons in 2014 International Semiconductor Laser Conference 1* (IEEE, 2014), 165–166. doi:10.1109/ISLC.2014.215.
227. Hurtado, A. & Javaloyes, J. Controllable spiking patterns in long-wavelength vertical cavity surface emitting lasers for neuromorphic photonics systems. *Applied Physics Letters* **107**, 241103. doi:10.1063/1.4937730 (2015).
228. Turconi, M. *et al.* Control of excitable pulses in an injection-locked semiconductor laser. *Physical Review E* **88**, 022923. doi:10.1103/PhysRevE.88.022923 (2013).
229. Garbin, B. *et al.* Refractory period of an excitable semiconductor laser with optical injection. *Physical Review E* **95**, 012214. doi:10.1103/PhysRevE.95.012214 (2017).
230. Lu, Y., Zhang, W., Fu, B. & He, Z. Frequency-switched photonic spiking neurons. *Optics Express* **30**, 21599. doi:10.1364/OE.456583 (2022).
231. Deng, T., Robertson, J. & Hurtado, A. Controlled Propagation of Spiking Dynamics in Vertical-Cavity Surface-Emitting Lasers: Towards Neuromorphic Photonic Networks. *IEEE Journal of Selected Topics in Quantum Electronics* **23**, 1–8. doi:10.1109/JSTQE.2017.2685140 (2017).
232. Deng, T. *et al.* Stable propagation of inhibited spiking dynamics in vertical-cavity surface-emitting lasers for neuromorphic photonic networks. *IEEE Access* **6**, 67951–67958. doi:10.1109/ACCESS.2018.2878940 (2018).
233. Robertson, J. *et al.* Toward Neuromorphic Photonic Networks of Ultrafast Spiking Laser Neurons. *IEEE Journal of Selected Topics in Quantum Electronics* **26**, 1–15. doi:10.1109/JSTQE.2019.2931215 (2020).
234. Xiang, S. *et al.* Cascadable Neuron-Like Spiking Dynamics in Coupled VCSELs Subject to Orthogonally Polarized Optical Pulse Injection. *IEEE Journal of Selected Topics in Quantum Electronics* **23**, 1–7. doi:10.1109/JSTQE.2017.2678170 (2017).

235. Robertson, J., Wade, E. & Hurtado, A. Electrically Controlled Neuron-Like Spiking Regimes in Vertical-Cavity Surface-Emitting Lasers at Ultrafast Rates. *IEEE Journal of Selected Topics in Quantum Electronics* **25**, 1–7. doi:10.1109/JSTQE.2019.2899040 (2019).
236. Robertson, J., Hejda, M., Bueno, J. & Hurtado, A. Ultrafast optical integration and pattern classification for neuromorphic photonics based on spiking VCSEL neurons. *Scientific Reports* **10**, 6098. doi:10.1038/s41598-020-62945-5 (2020).
237. Zhang, Y. *et al.* All-optical neuromorphic binary convolution with a spiking VCSEL neuron for image gradient magnitudes. *Photonics Research* **9**, B201. doi:10.1364/PRJ.412141 (2021).
238. Robertson, J. *et al.* Image edge detection with a photonic spiking VCSEL-neuron. *Optics Express* **28**, 37526. doi:10.1364/OE.408747 (2020).
239. Robertson, J. *et al.* Ultrafast neuromorphic photonic image processing with a VCSEL neuron. *Scientific Reports* **12**, 4874. doi:10.1038/s41598-022-08703-1 (2022).
240. Owen-Newns, D., Robertson, J., Hejda, M. & Hurtado, A. GHz Rate Neuromorphic Photonic Spiking Neural Network With a Single Vertical-Cavity Surface-Emitting Laser (VCSEL). *IEEE Journal of Selected Topics in Quantum Electronics* **29**, 1–10. doi:10.1109/JSTQE.2022.3205716. arXiv: 2208.06347 (2023).
241. Chen, Z. *et al.* Deep learning with coherent VCSEL neural networks. *Nature Photonics* **17**, 723–730. doi:10.1038/s41566-023-01233-w (2023).
242. Skalli, A. *et al.* Photonic neuromorphic computing using vertical cavity semiconductor lasers. *Optical Materials Express* **12**, 2395. doi:10.1364/OME.450926 (2022).
243. Prucnal, P. R. & Shastri, B. J. *Neuromorphic Photonics* 1st Editin. 440 pp. (CRC Press, Boca Raton, 2017).
244. Peng, H.-T. *et al.* Temporal Information Processing With an Integrated Laser Neuron. *IEEE Journal of Selected Topics in Quantum Electronics* **26**, 1–9. doi:10.1109/JSTQE.2019.2927582 (2020).
245. Nahmias, M. A., Shastri, B. J., Tait, A. N. & Prucnal, P. R. A Leaky Integrate-and-Fire Laser Neuron for Ultrafast Cognitive Computing. *IEEE Journal of Selected Topics in Quantum Electronics* **19**, 1–12. doi:10.1109/JSTQE.2013.2257700 (2013).
246. Han, Y. *et al.* Spiking dynamics and synchronization properties of optical neurons based on VCSEL-SAs. *Nonlinear Dynamics*. doi:10.1007/s11071-021-06699-3 (2021).
247. Song, Z. *et al.* Spike Sequence Learning in a Photonic Spiking Neural Network Consisting of VCSELs-SA With Supervised Training. *IEEE Journal of Selected Topics in Quantum Electronics* **26**, 1–9. doi:10.1109/JSTQE.2020.2975564 (2020).

248. Xiang, S. *et al.* Computing Primitive of Fully VCSEL-Based All-Optical Spiking Neural Network for Supervised Learning and Pattern Classification. *IEEE Transactions on Neural Networks and Learning Systems* **32**, 2494–2505. doi:10.1109/TNNLS.2020.3006263 (2020).
249. Zhang, Y. *et al.* The winner-take-all mechanism for all-optical systems of pattern recognition and max-pooling operation. *Journal of Lightwave Technology* **38**, 5071–5077. doi:10.1109/JLT.2020.3000670 (2020).
250. Gao, S. *et al.* Motion detection and direction recognition in a photonic spiking neural network consisting of VCSELs-SA. *Optics Express* **30**, 31701. doi:10.1364/OE.465653 (2022).
251. Song, Z. *et al.* Photonic spiking neural network based on excitable VCSELs-SA for sound azimuth detection. *Optics Express* **28**, 1561. doi:10.1364/OE.381229 (2020).
252. Toomey, J. P. *et al.* Stability of the nonlinear dynamics of an optically injected VCSEL. *Optics Express* **20**, 10256. doi:10.1364/OE.20.010256 (2012).
253. Al-Seyab, R. *et al.* Dynamics of VCSELs Subject to Optical Injection of Arbitrary Polarization. *IEEE Journal of Selected Topics in Quantum Electronics* **19**, 1700512–1700512. doi:10.1109/JSTQE.2013.2239614 (2013).
254. Hejda, M. *et al.* Neuromorphic encoding of image pixel data into rate-coded optical spike trains with a photonic VCSEL-neuron. *APL Photonics* **6**, 060802. doi:10.1063/5.0048674 (2021).
255. Hejda, M., Robertson, J., Bueno, J. & Hurtado, A. Spike-based information encoding in vertical cavity surface emitting lasers for neuromorphic photonic systems. *Journal of Physics: Photonics* **2**, 044001. doi:10.1088/2515-7647/aba670 (2020).
256. Selmi, F. *et al.* Spike latency and response properties of an excitable micropillar laser. *Physical Review E* **94**, 042219. doi:10.1103/PhysRevE.94.042219 (2016).
257. Gollisch, T. & Meister, M. Rapid Neural Coding in the Retina with Relative Spike Latencies. *Science* **319**, 1108–1111. doi:10.1126/science.1149639 (2008).
258. Zhang, Z. *et al.* Controllable spiking dynamics in cascaded VCSEL-SA photonic neurons. *Nonlinear Dynamics*. doi:10.1007/s11071-019-05339-1 (2019).
259. Robertson, J., Deng, T., Javaloyes, J. & Hurtado, A. Controlled inhibition of spiking dynamics in VCSELs for neuromorphic photonics: theory and experiments. *Optics Letters* **42**, 1560. doi:10.1364/ol.42.001560 (2017).
260. Heuser, T. *et al.* Developing a photonic hardware platform for brain-inspired computing based on 5×5 VCSEL arrays. *Journal of Physics: Photonics* **2**, 044002. doi:10.1088/2515-7647/aba671 (2020).
261. Gu, M. *et al.* Perspective on 3D vertically-integrated photonic neural networks based on VCSEL arrays. *Nanophotonics* **0**. doi:10.1515/nanoph-2022-0437 (2023).

262. Gerstner, W. & Kistler, W. M. *Spiking Neuron Models: Single Neurons, Populations, Plasticity* doi:10.1017/CB09780511815706 (Cambridge University Press, Cambridge, 2002).
263. Hejda, M. *et al.* *Neuromorphic rate-coded representation of digital image data with a VCSEL-neuron* in. European Semiconductor Laser Workshop (ESLW2020) (Technical University of Eindhoven, Eindhoven, NL, 2020).
264. Hejda, M. *et al.* *Interfacing spiking VCSEL-neurons with silicon photonics weight banks towards integrated neuromorphic photonic systems* 2023.
265. Tait, A. N. *et al.* Neuromorphic photonic networks using silicon photonic weight banks. *Scientific Reports* **7**, 7430. doi:10.1038/s41598-017-07754-z (2017).
266. Ironside, C. N., Romeira, B. & Figueiredo, J. M. L. *Resonant Tunneling Diode Photonics: Devices and Applications* doi:10.1088/2053-2571/ab3a9a (Morgan & Claypool Publishers, 2019).
267. Hejda, M. *et al.* Artificial optoelectronic spiking neuron based on a resonant tunnelling diode coupled to a vertical cavity surface emitting laser. *Nanophotonics* **12**, 857–867. doi:10.1515/nanoph-2022-0362 (2022).
268. Sollner, T. C. L. G. *et al.* Resonant tunneling through quantum wells at frequencies up to 2.5 THz. *Applied Physics Letters* **43**, 588–590. doi:10.1063/1.94434 (1983).
269. Dessouki, A. A. S., Abdallah, R. M. & Aly, M. H. A Simplified Analytical Technique for High Frequency Characterization of Resonant Tunneling Diode. *Advances in Electrical and Computer Engineering* **14**, 87–94. doi:10.4316/AECE.2014.04013 (2014).
270. Asada, M. & Suzuki, S. *Compact THz oscillators with resonant tunneling diodes and application to high-capacity wireless communications* in *ICECom 2013* ICECom 2013 (2013), 1–5. doi:10.1109/ICECom.2013.6684744.
271. Cimbri, D., Wang, J., Al-Khalidi, A. & Wasige, E. Resonant Tunneling Diodes High-Speed Terahertz Wireless Communications - A Review. *IEEE Transactions on Terahertz Science and Technology*, 1–1. doi:10.1109/TTHZ.2022.3142965 (2022).
272. Romeira, B., Seunarine, K., Ironside, C. N. & Kelly, A. E. A Self-Synchronized Optoelectronic Oscillator Based on an RTD Photodetector and a Laser Diode. *IEEE Photonics Technology Letters* **23**, 3. doi:10.1109/LPT.2011.2154320 (2011).
273. Blakesley, J. C. *et al.* Efficient Single Photon Detection by Quantum Dot Resonant Tunneling Diodes. *Physical Review Letters* **94**, 067401. doi:10.1103/PhysRevLett.94.067401 (2005).
274. Pfenning, A. *et al.* Sensitivity of resonant tunneling diode photodetectors. *Nanotechnology* **27**, 355202. doi:10.1088/0957-4484/27/35/355202 (2016).
275. Li, H. W. *et al.* Quantum dot resonant tunneling diode for telecommunication wavelength single photon detection. *Applied Physics Letters* **91**, 073516. doi:10.1063/1.2768884 (2007).

276. Weng, Q. *et al.* Quantum dot single-photon switches of resonant tunneling current for discriminating-photon-number detection. *Scientific Reports* **5**, 9389. doi:10.1038/srep09389 (2015).
277. Pfenning, A. *et al.* Single-Photon Counting with Semiconductor Resonant Tunneling Devices. *Nanomaterials* **12**, 2358. doi:10.3390/nano12142358 (2022).
278. Miyamoto, T., Yamaguchi, A. & Mukai, T. Terahertz imaging system with resonant tunneling diodes. *Japanese Journal of Applied Physics* **55**, 032201. doi:10.7567/JJAP.55.032201 (2016).
279. Hartmann, F. *et al.* Nanowatt logic stochastic resonance in branched resonant tunneling diodes. *Applied Physics Letters* **98**, 032110. doi:10.1063/1.3548539 (2011).
280. Nishida, Y. *et al.* Terahertz coherent receiver using a single resonant tunnelling diode. *Scientific Reports* **9**, 18125. doi:10.1038/s41598-019-54627-8 (2019).
281. Al-Taai, Q. *et al.* Towards an Excitable Microwave Spike Generator for Future Neuromorphic Computing in 2021 16th European Microwave Integrated Circuits Conference (EuMIC) 2021 16th European Microwave Integrated Circuits Conference (EuMIC) (2022), 386–389. doi:10.23919/EuMIC50153.2022.9783686.
282. Malysheva, E. *et al.* Impact of High Temperature Post-Treatment on Photoluminescence Performance of Passivated InP/In_{0.53}Ga_{0.47}As/InP Nanopillars in 2021 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC) 2021 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC) (IEEE, Munich, Germany, 2021), 1–1. doi:10.1109/CLEO/Europe-EQEC52157.2021.9541551.
283. Van Mai, T., Suzuki, Y., Suzuki, S. & Asada, M. Novel RTD Oscillator with Simplified Structure and Fabrication Process in 2020 45th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz) 2020 45th International Conference on Infrared, Millimeter and Terahertz Waves (IRMMW-THz) (IEEE, Buffalo, NY, USA, 2020), 1–2. doi:10.1109/IRMMW-THz46771.2020.9370941.
284. Figueiredo, J. M. L. *Optoelectronic Properties of Resonant Tunnelling Diodes* Doctoral thesis (Universidade do Porto, 2000). 260 pp.
285. Sun, Z. Z. *et al.* Self-sustained current oscillations in superlattices and the van der Pol equation. *Applied Physics Letters* **87**, 182110. doi:10.1063/1.2126149 (2005).
286. Ginoux, J.-M. & Letellier, C. Van der Pol and the history of relaxation oscillations: Toward the emergence of a concept. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **22**, 023120. doi:10.1063/1.3670008 (2012).
287. Schulman, J. N., De Los Santos, H. J. & Choi, D. H. Physics-based RTD current-voltage equation. *IEEE Electron Device Letters* **17**, 220–222. doi:10.1109/55.491835 (1996).

288. Slight, T. J. *et al.* A Liénard oscillator resonant tunnelling diode-laser diode hybrid integrated circuit: Model and experiment. *IEEE Journal of Quantum Electronics* **44**, 1158–1163. doi:10.1109/JQE.2008.2000924 (2008).
289. Romeira, B., Figueiredo, J. M. L. & Javaloyes, J. Delay dynamics of neuromorphic optoelectronic nanoscale resonators: Perspectives and applications. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **27**, 114323. doi:10.1063/1.5008888 (2017).
290. Rowat, P. F. & Selverston, A. I. Modeling the gastric mill central pattern generator of the lobster with a relaxation-oscillator network. *Journal of Neurophysiology* **70**, 24. doi:10.1152/jn.1993.70.3.1030 (1993).
291. Feiginov, M. Frequency Limitations of Resonant-Tunnelling Diodes in Sub-THz and THz Oscillators and Detectors. *Journal of Infrared, Millimeter, and Terahertz Waves* **40**, 365–394. doi:10.1007/s10762-019-00573-5 (2019).
292. Kim, C. & Brandli, A. High-Frequency High-Power Operation of Tunnel Diodes. *IRE Transactions on Circuit Theory* **8**, 416–425. doi:10.1109/TCT.1961.1086849 (1961).
293. Figueiredo, J. M. L., Romeira, B., Slight, T. & Ironside, C. N. in *Advances in Optical and Photonic Devices* (ed Kim, K. Y.) 1st (InTech, 2010). doi:10.5772/7146.
294. Ortega-Piwonka, I. *et al.* Bursting and Excitability in Neuromorphic Resonant Tunneling Diodes. *Physical Review Applied* **15**, 034017. doi:10.1103/PhysRevApplied.15.034017 (2021).
295. Liu, Q., Seabaugh, A., Chahal, P. & Morris, F. J. Unified AC model for the resonant tunneling diode. *IEEE Transactions on Electron Devices* **51**, 653–657. doi:10.1109/TED.2004.825795 (2004).
296. Sheng, H., Chua, S.-J. & Sinkkonen, J. Capacitance property of resonant tunnelling diodes. *Modern Physics Letters B*. doi:10.1142/S0217984993000461 (1993).
297. Morariu, R. *et al.* Accurate Small-Signal Equivalent Circuit Modeling of Resonant Tunneling Diodes to 110 GHz. *IEEE Transactions on Microwave Theory and Techniques* **67**, 4332–4340. doi:10.1109/TMTT.2019.2939321 (2019).
298. Cornescu, A. C. *et al.* High-Efficiency Bias Stabilization for Resonant Tunneling Diode Oscillators. *IEEE Transactions on Microwave Theory and Techniques* **67**, 3449–3454. doi:10.1109/TMTT.2019.2916602 (2019).
299. Wang, L., Figueiredo, J. M. L., Ironside, C. N. & Wasige, E. DC Characterization of Tunnel Diodes Under Stable Non-Oscillatory Circuit Conditions. *IEEE Transactions on Electron Devices* **58**, 343–347. doi:10.1109/TED.2010.2091507 (2011).
300. Romeira, B. *et al.* Excitability and optical pulse generation in semiconductor lasers driven by resonant tunneling diode photo-detectors. *Optics Express* **21**, 20931. doi:10.1364/OE.21.020931 (2013).

301. Bao, M. & Wang, K. L. Accurately measuring current-voltage characteristics of tunnel diodes. *IEEE Transactions on Electron Devices* **53**, 2564–2568. doi:10.1109/TED.2006.882281 (2006).
302. Cantú, H. I. *et al.* Resonant Tunneling Diode Optoelectronic Circuits Applications in Radio-Over-Fiber Networks. *IEEE Transactions on Microwave Theory and Techniques* **60**, 2903–2912. doi:10.1109/TMTT.2012.2206606 (2012).
303. Romeira, B. *et al.* Nonlinear Dynamics of Resonant Tunneling Optoelectronic Circuits for Wireless/Optical Interfaces. *IEEE Journal of Quantum Electronics* **45**, 1436–1445. doi:10.1109/JQE.2009.2028084 (2009).
304. Kingston, S. L. & Thamilmaran, K. Bursting Oscillations and Mixed-Mode Oscillations in Driven Liénard System. *International Journal of Bifurcation and Chaos* **27**, 1730025. doi:10.1142/S0218127417300257 (2017).
305. Hartmann, F. *et al.* Light-induced stochastic resonance in a nanoscale resonant-tunneling diode. *Applied Physics Letters* **98**, 96–99. doi:10.1063/1.3600329 (2011).
306. Romeira, B. *et al.* Photo-Detectors Integrated with Resonant Tunneling Diodes. *Sensors* **13**, 9464–9482. doi:10.3390/s130709464 (2013).
307. Moise, T. S. & Kao, Y.-C. Optically switched resonant tunneling diodes. *Applied Physics Letters* **66**, 1104–1106. doi:10.1063/1.113826 (1995).
308. Coêlho, I. J. S., Martins-Filho, J. F., Figueiredo, J. M. L. & Ironside, C. N. Modeling of light-sensitive resonant-tunneling-diode devices. *Journal of Applied Physics* **95**, 8258–8263. doi:10.1063/1.1728290 (2004).
309. Yokoyama, H. & Brorson, S. D. Rate equation analysis of microcavity lasers. *Journal of Applied Physics* **66**, 4801–4805. doi:10.1063/1.343793 (1989).
310. Mork, J. & Lippi, G. L. Rate equation description of quantum noise in nanolasers with few emitters. *Applied Physics Letters* **112**, 141103. doi:10.1063/1.5022958 (2018).
311. Bjork, G. & Yamamoto, Y. Analysis of semiconductor microcavity lasers using rate equations. *IEEE Journal of Quantum Electronics* **27**, 2386–2396. doi:10.1109/3.100877 (1991).
312. Rice, P. R. & Carmichael, H. J. Photon statistics of a cavity-QED laser: A comment on the laser phase-transition analogy. *Physical Review A* **50**, 4318–4329. doi:10.1103/PhysRevA.50.4318 (1994).
313. Ortega-Piwonka, I. *et al.* Spike propagation in a nanolaser-based optoelectronic neuron. *Optical Materials Express*. doi:10.1364/OME.451706 (2022).
314. Mørk, J. & Yvind, K. Squeezing of intensity noise in nanolasers and nanoLEDs with extreme dielectric confinement. *Optica* **7**, 1641. doi:10.1364/OPTICA.402190 (2020).
315. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* **17**, 261–272. doi:10.1038/s41592-019-0686-2 (2020).

316. Lam, S. K., Pitrou, A. & Seibert, S. *Numba: a LLVM-based Python JIT compiler* in *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC* (Association for Computing Machinery, New York, NY, USA, 2015), 1–6. doi:10.1145/2833157.2833162.
317. Romeira, B. *et al.* Regenerative memory in time-delayed neuromorphic photonic resonators. *Scientific Reports* **6**, 19510. doi:10.1038/srep19510. arXiv: 1503.07781 (2016).
318. Selmi, F. *et al.* Relative Refractory Period in an Excitable Semiconductor Laser. *Physical Review Letters* **112**, 183902. doi:10.1103/PhysRevLett.112.183902 (2014).
319. Al-Taai, Q. R. A. *et al.* Optically-triggered deterministic spiking regimes in nanostructure resonant tunnelling diode-photodetectors. *Neuromorphic Computing and Engineering* **3**, 034012. doi:10.1088/2634-4386/acf609 (2023).
320. Hejda, M. *et al.* Resonant Tunneling Diode Nano-Optoelectronic Excitable Nodes for Neuromorphic Spike-Based Information Processing. *Physical Review Applied* **17**, 024072. doi:10.1103/PhysRevApplied.17.024072 (2022).
321. Wang, W. *et al.* Learning of spatiotemporal patterns in a spiking neural network with resistive switching synapses. *Science Advances* **4**, eaat4752. doi:10.1126/sciadv.aat4752 (2018).
322. Stark, P. *et al.* Opportunities for integrated photonic neural networks. *Nanophotonics* **9**, 4221–4232. doi:10.1515/nanoph-2020-0297 (2020).
323. Agmon-Snir, H., Carr, C. E. & Rinzel, J. The role of dendrites in auditory coincidence detection. *Nature* **393**, 268–272. doi:10.1038/30505 (1998).
324. Stevens, C. F. & Zador, A. M. Input synchrony and the irregular firing of cortical neurons. *Nature Neuroscience* **1**, 210–217. doi:10.1038/659 (1998).
325. George, J. K., Soci, C., Miscuglio, M. & Sorger, V. J. Symmetry perception with spiking neural networks. *Scientific Reports* **11**, 5776. doi:10.1038/s41598-021-85232-3 (2021).
326. Pammi, V. A., Alfaro-Bittner, K., Clerc, M. G. & Barbay, S. Photonic Computing With Single and Coupled Spiking Micropillar Lasers. *IEEE Journal of Selected Topics in Quantum Electronics* **26**, 1–7. doi:10.1109/JSTQE.2019.2929187 (2020).
327. Peng, H.-T. *et al.* Neuromorphic Photonic Integrated Circuits. *IEEE Journal of Selected Topics in Quantum Electronics* **24**, 1–15. doi:10.1109/JSTQE.2018.2840448 (2018).
328. Shastri, B. J. *et al.* Spike processing with a graphene excitable laser. *Scientific Reports* **6**, 19126. doi:10.1038/srep19126 (2016).
329. Nahmias, M. A. *et al.* *A Laser Spiking Neuron in a Photonic Integrated Circuit* 2020. arXiv: 2012.08516.
330. Mekemeza-Ona, K., Routier, B. & Charbonnier, B. All optical Q-switched laser based spiking neuron. *Frontiers in Physics* **10**, 1017714. doi:10.3389/fphy.2022.1017714 (2022).

331. Hejda, M. *et al.* Spiking Behaviour in Laterally-Coupled Pairs of VCSELs With Applications in Neuromorphic Photonics. *IEEE Journal of Selected Topics in Quantum Electronics*, 1–10. doi:10.1109/JSTQE.2022.3218950 (2022).
332. Romeira, B. *et al.* Stochastic induced dynamics in neuromorphic optoelectronic oscillators. *Optical and Quantum Electronics* **46**, 1391–1396. doi:10.1007/s11082-014-9905-3 (2014).
333. Zhang, W. *et al.* Tuneable presynaptic weighting in optoelectronic spiking neurons built with laser-coupled resonant tunneling diodes. *Journal of Physics D: Applied Physics* **56**. doi:10.1088/1361-6463/aca914 (2022).
334. Mankins, J. C. Technology readiness assessments: A retrospective. *Acta Astronautica* **65**, 1216–1223. doi:10.1016/j.actaastro.2009.03.058 (2009).
335. Hill, M. T. & Gather, M. C. Advances in small lasers. *Nature Photonics* **8**, 908–918. doi:10.1038/nphoton.2014.239 (2014).
336. Church, S. A. *et al.* Holistic Nanowire Laser Characterization as a Route to Optimal Design. *Advanced Optical Materials*, 2202476. doi:10.1002/adom.202202476 (2023).
337. Reimer, M. E. *et al.* Bright single-photon sources in bottom-up tailored nanowires. *Nature Communications* **3**, 737. doi:10.1038/ncomms1746 (2012).
338. Adão, R. M. R. *et al.* Two-photon polymerization simulation and fabrication of 3D microprinted suspended waveguides for on-chip optical interconnects. *Optics Express* **30**, 9623. doi:10.1364/OE.449641 (2022).
339. Grabulosa, A. *et al.* $(3+1)D$ -printed adiabatic 1-to- N broadband couplers 2022. arXiv: 2208.09767[physics].
340. Grabulosa, A. *et al.* Additive 3D photonic integration that is CMOS compatible 2023. doi:10.48550/arXiv.2301.00983. arXiv: 2301.00983[physics].
341. Chiles, J. *et al.* Design, fabrication, and metrology of 10×100 multi-planar integrated photonic routing manifolds for neural networks. *APL Photonics* **3**, 106101. doi:10.1063/1.5039641 (2018).
342. Nuck, M. *et al.* 3D photonic integrated 4×4 multi-mode interference coupler in *Integrated Optics: Devices, Materials, and Technologies XXIII* Integrated Optics: Devices, Materials, and Technologies XXIII (eds García-Blanco, S. M. & Cheben, P.) (SPIE, San Francisco, United States, 2019), 47. doi:10.1117/12.2509776.
343. Yik, J. *et al.* NeuroBench: Advancing Neuromorphic Computing through Collaborative, Fair and Representative Benchmarking 2023. arXiv: 2304.04640[cs].
344. Simons, T. & Lee, D.-J. A Review of Binarized Neural Networks. *Electronics* **8**, 661. doi:10.3390/electronics8060661 (2019).

345. Courbariaux, M. *et al.* *Binarized Neural Networks: Training Deep Neural Networks with Weights and Activations Constrained to +1 or -1* 2016. arXiv: 1602.02830.
346. Kendall, J. *et al.* *Training End-to-End Analog Neural Networks with Equilibrium Propagation* 2020. doi:10.48550/arXiv.2006.01981. arXiv: 2006.01981[cs].
347. Hinton, G. *The Forward-Forward Algorithm: Some Preliminary Investigations* 2022.
348. Murmann, B. *Energy limits in A/D converters* in *2013 IEEE Faible Tension Faible Consommation 2013 IEEE Faible Tension Faible Consommation (FTFC)* (IEEE, Paris, France, 2013), 1–4. doi:10.1109/FTFC.2013.6577781.
349. Alanis, J. A., Robertson, J., Hejda, M. & Hurtado, A. Weight Adjustable Photonic Synapse by Non-Linear Gain in a Vertical Cavity Semiconductor Optical Amplifier. *Applied Physics Letters*. doi:10.1063/5.0064374 (2021).
350. Robertson, J., Alanis, J. A., Hejda, M. & Hurtado, A. Photonic synaptic system for MAC operations by interconnected vertical cavity surface emitting lasers. *Optical Materials Express* **12**, 1417–1426. doi:10.1364/OME.450923 (2022).
351. Haegele, S. *et al.* Large field-of-view holographic imager with ultra-high phase sensitivity using multi-angle illumination. *Optics and Lasers in Engineering* **161**, 107315. doi:10.1016/j.optlaseng.2022.107315 (2023).
352. Owen-Newns, D., Robertson, J., Hejda, M. & Hurtado, A. Photonic Spiking Neural Networks with Highly Efficient Training Protocols for Ultrafast Neuromorphic Computing Systems. *Intelligent Computing* **0**. doi:10.34133/icomputing.0031 (ja 2023).