Photonic computing with coupled spiking micropillars and extreme event prediction in microcavity lasers

Venkata Anirudh Pammi

université PARIS-SACLAY

Photonic computing with coupled spiking micropillars and extreme event prediction in microcavity lasers

Calcul photonique avec des micropiliers excitables couplés et prédiction d'événements extrêmes dans des microcavités lasers

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 572 : ondes et matière (EDOM)
Spécialité de doctorat : Physique
Unité de recherche : Université Paris-Saclay, CNRS, Centre de Nanosciences et de Nanotechnologies, 91120, Palaiseau, France
Graduate School : Physique, Référent : Faculté des sciences d'Orsay

Thèse présentée et soutenue à Paris-Saclay, le 20 Septembre 2021, par

Venkata Anirudh PAMMI

Composition du jury :

Massimo GIUDICI Professeur, Université Côte d'Azur	President
Daniel BRUNNER Chargé de recherche CNRS, HDR, Femto-ST	Rapporteur & Examinateur
Marc SCIAMANNA Professeur, CentraleSupélec - Campus de Metz	Rapporteur & Examinateur
Peter BIENSTMAN Professeur, Université de Gand	Examinateur
Alice MIZRAHI Ingénieure de recherche, Thales Research & Technology, Unité Mixte de Physique CNRS/Thales	Examinatrice
Sylvain BARBAY Directeur de recherche CNRS, C2N, Université Paris-Saclay	Directeur de thèse

Thèse de doctorat

NNT: 2021UPASP080

Dedicated to my parents, who have been a constant support throughout and put in the effort to develop a working understanding of my various research projects.

Acknowledgements

I would first like to thank my parents, who have essentially made me the person I am today through their constant guidance and support. A PhD programme is a journey of both professional and personal growth; a journey that I could not have undertaken without the support of certain people, whom I must thank.

I am very grateful to my thesis director, Sylvain Barbay, for being a patient and helpful teacher and for answering my questions, physics or otherwise. His guidance and help were instrumental in this thesis. I would then like to thank our collaborators: Soizic Terrien, Neil G.R. Broderick, Bernd Krauskopf, Marcel G. Clerc, Saliya Coulibaly, Rémy Braive, Karin Alfaro-Bittner, Gregorio Gonzalez-Cortes, Isabelle Sagnes, Abdelmounaim Harouri, Guy Aubin, and Nathalie Bardou for their theoretical and experimental contributions to this thesis. I would like to thank my colleagues in the TONIQ group for adding a pleasant social aspect to my work life. I also enjoyed my overlap working with Amir Masominia and Léna Soun on various neuromorphic and cleanroom projects.

For my personal growth, I would first like to thank my friends at the *Maison de l'Inde*, a student residence at *Cité internationale universitaire de Paris* for all their help, conversations and support during my thesis; their involvement made my stay in Île-de-France very memorable! I would also like to mention my sincere gratitude to my friends from India, who in their many ways helped me and continue to be an integral part of my life.

Lastly, I would like to express my gratitude to my mentors, coordinators, and peers from my Master's programme, EuroPhotonics, for supporting me through this journey.

Contents

List of Figures

Chapter 1

Introduction

Understanding neuronal computational circuits and mimicking them has been a long-standing endeavour in many fields, especially since the brain is capable of representing information it receives from sensory inputs and acting on them with extraordinary efficiency. However, predetermined algorithms are often not very suitable to represent and act upon such data; sometimes the algorithms need to be developed from the very data presented to them. This is the foundational idea behind the field of artificial intelligence (AI).

A preliminary understanding of the structure of the nervous system was brought about by scientists such as Camillo Golgi and Santiago Ramón y Cajal, who were awarded the Nobel Prize in Physiology in 1906. Though they shared the Nobel prize, the two scientists believed in different schools of thought about the nervous structure. Camillo Golgi believed in the reticular theory to describe the nervous system, postulating that the nervous system is a single continuous network, whereas Santiago Ramón y Cajal believed in the neuron theory, which postulated that the nervous system is made up of single discrete units (neurons) connected through synapses. Inspired by the neuron theory in 1943, two engineers Warren McCullogh and Walter Pitts proposed the first mathematical model for the neuron. This initial model represented neurons as binary entities with inputs (excitatory and inhibitory) and a threshold for producing an output. Building on that, Rosenblatt in 1958 proposed a computational model called the Perceptron [1], which arguably was the world's first artificial neural network. The Perceptron marked a significant improve-

ment as it included learning rules and showed promising results on some tasks. It is worth noting that the initially proposed Perceptron had one layer of neurons and could only learn linearly separable patterns. This severe constraint was not commensurate with the hype and led to a decline in interest surrounding Perceptrons. It was several years before it was discovered that multilayer Perceptrons could learn nonlinear decision boundaries and possessed much greater computational power. Scientific funding or lack thereof is often a result of perception surrounding a subject; the bad reputation of single layer Perceptrons probably was one of the reasons that caused a decline in funding for AI research for several following years.

While several refinements were made to the Perceptron algorithm over the years, one of the key contributions to the field was made in 1986, when researchers showed that an algorithm called backpropagation could be used to teach representations to networks of neuron like units [2][1].

Although the work of Alan Hodgkin, and Andrew Huxley (awarded the Nobel Prize in Physiology or Medicine in 1963) on ion channels and excitable neurons [3, 4] predated the Perceptron model, these findings were not a subject of extensive study by the AI and machine learning community for a long time. One can view this as a divergence in the paths of the neuroscience and the AI communities.

From an engineering perspective, the AI community had learnt that coupling discrete nonlinear units using weights could give rise to networks with an apparent semblance of intelligence. Armed with an algorithm to train such networks, researchers began their extensive quest on training models for increasingly complicated tasks. Some of the major limitations before the turn of the 21^{st} century were the lack of compute power and data available to scale up such networks; these limitations were eventually removed with the semiconductor and digital revolution. Subsequently, neural networks gained even more traction. An alternative framework was introduced in 2002 [5] to model neural computations which relied on several coupled spiking neurons. Independently, an equivalent model called echo state networks were introduced [6] which essentially had a similar

[1] It is important to note that this paper is often incorrectly credited with inventing the backpropagation algorithm. This paper only claims to be a clear demonstration showing that the algorithm can be used to teach representations. Several researchers independently invented similar algorithms.

structure but the definition of a neuron was more abstract and essentially any nonlinearity could be used to model a neuron. These two models can be broadly combined into a single class called reservoir computing. The main power of this technique arises from the complex dynamics of coupled nonlinear nodes.

At this point we would like to insist on the difference between neural networks and neuronal networks. The former refers to the engineering abstraction comprising of nonlinear coupled nodes and the latter refers to the study of biological networks of neurons. The study of neuronal networks is an extremely complicated endeavour as one often does not have access to a connectome but even in cases where a connectome is available (for small organisms like C. Elegans [7]) the mechanisms for computing are not trivial. Over the years, several neural network models have shown remarkable success at several tasks. One such feat was when a convolutional neural network (ref. [8]) in 2012 won the ImageNet Large Scale Visual Recognition Challenge. This feat was one of the reasons for the renewed interest in the field of neural networks.

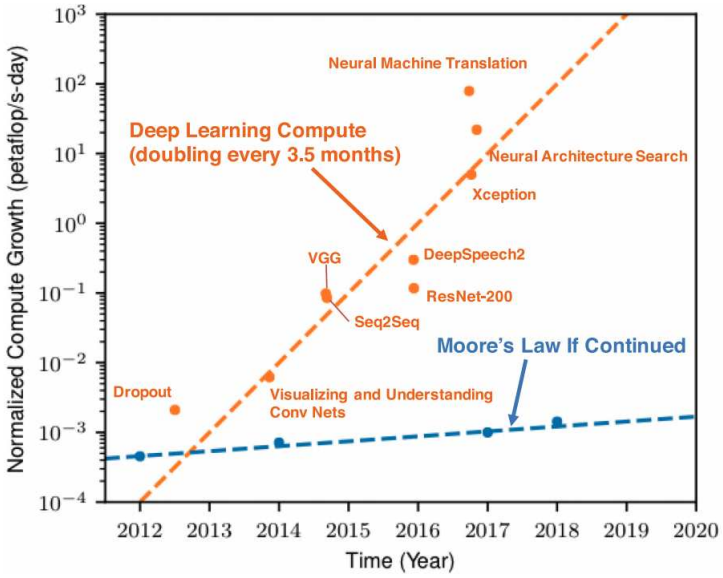


Figure 1.1: The disparity between the compute power required to train state-of-the-art neural networks and that supplied by Moore's law

The trend in orange shows the amount of compute power required to train the state-of-the-art algorithms (normalized to petaflop/s-day). The trend in blue shows the compute growth predicted by Moore's law. A petaflop/s-day corresponds to 10^{15} operations per second for one day. This amounts to a total of 8.64×10^{19} operations $(10^{15} * 3600 * 24)$. Figure adapted from [9]

While the notion of neural networks was inspired from experimental findings in neuroscience, today, the two fields have diverged significantly. Neural networks in their most general form are powerful optimization algorithms. Over the years, neural networks have been trained for increasingly complicated tasks and this is reflected in their power requirement. Figure 1.1 plots the growth in the complexity of modern neural network algorithms and the projected growth of compute power offered by electronics. The computing power demanded by the algorithms is growing at a much faster rate and cannot be supported by electronics. One of the most popular examples is the comparison of energy consumption between Google's AlphaGo program and the human brain playing a game of Go. It is estimated that the human brain consumes approximately 10^6 times less energy as compared to Google's state of the art AI platform. This is a pertinent comparison as it is a popular opinion within the community that reinforcement learning, the work horse behind AlphaGo, is a plausible path for general intelligence.

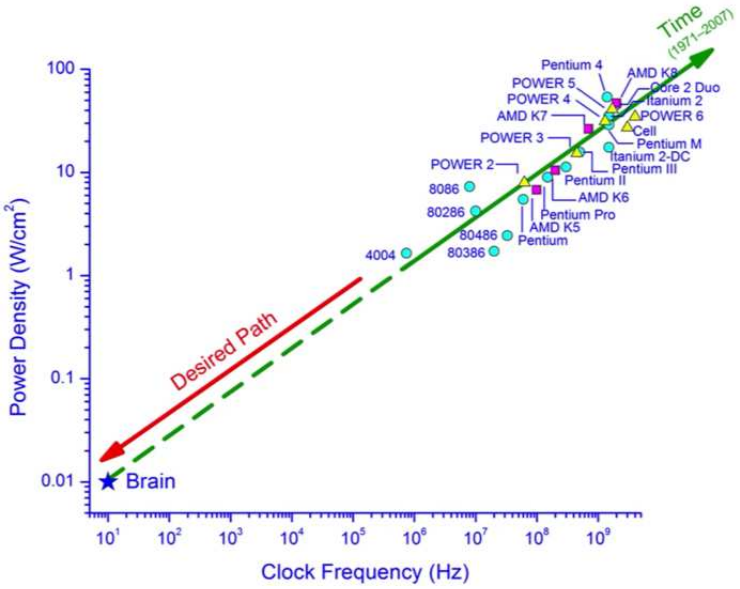


Figure 1.2: Power density and clock frequency of microprocessors and the brain
Source: IBM Research-Almaden Cognitive Systems Colloquium, 2014

This raises important and fundamental questions about the architecture of modern neural networks and their implementation on conventional platforms. Figure 1.2 depicts a trend (in power density and clock frequency) observed in microprocessors that is diametrically opposite to that of the brain. This shows that while the architecture of neural networks is

inspired from biology, its current physical implementation is on drastically different primitives. Further, experimental findings show that the brain possesses an extensively parallel network of neurons that is absent in traditional electronic implementations. At this point we can raise two important questions:

- Can we ameliorate the hardware implementation of neural networks?
- Can we use architectures inspired from biology to achieve energy efficient computing?

There are several interesting electronic implementations that can answer both these questions; some currently available commercial solutions include the development of Graphics Processing Units and Tensor Processing Units. Other electronic technologies such as memristors [10], Resistive RAM based binarized neural networks [11], and spintronic oscillators [12] have also answered the above mentioned questions. A general review on the physics of neuromorphic computing can be found in ref. [13].t

Despite the maturity of electronics and their relative flexibility to implement complex circuits, optical and optoelectronic technologies can offer significant advantages in terms of parallelism [14], small device footprint, and energy efficiency over electronics in certain implementations [15, 16]. In the following paragraphs, we will present a non-exhaustive list of technologies with an emphasis on optical implementations.

At the heart of neural network training and inference are repeated matrix vector products (or multiply and accumulate operations (MAC)). One of the solutions proposed to perform arbitrary matrix vector products in the optical domain is shown in ref. [17]. While this architecture provides a path for direct hardware implementation of neural networks, may have limitations in scaling up as the number of components required scale quadratically with the size of the matrix. In ref. [18] an integrated photonic accelerator was demonstrated to have the capability of performing 10^{12} MAC/s. This circuit was realized with the help of microring resonators and phase change materials (PCMs). In ref. [19] presents an analysis on the different methods to perform such MAC operations and their relative efficiencies.

Another direction of research is to develop new computing primitives and architectures

that mimic the behaviour of neuronal systems. This category encompasses a large range of architectures such as:

- In memory computing: In an oversimplified way, the memory and computational units in the brain (synaptic weights and soma respectively) are located in close proximity as opposed to von Neumann architectures where the two units are separate and a significant portion of the energy is spent on moving the data between the two units. PCMs can be used to emulate this proximity and demonstrate optical synapses [20] and spiking neural networks [21].

- Coupled nonlinear nodes: Evidence from neuroscience shows that the dynamics of several coupled neurons give rise to intelligence. While the exact mechanisms are not clear, having a coupled neurons seems to be essential. This was the foundational idea behind neural networks and reservoir computing. Reservoir computing provides additional flexibility as it is designed to operate with essentially random weights between nodes as opposed to requiring precise trainable weights. Early numerical illustrations in the optical domain used a semiconductor optical amplifier as the nonlinear node [22, 23]. Following this, several architectures for reservoir computing have been proposed using all-optical systems [24], silicon photonic chips [25], optoelectronic systems [26, 27], semiconductor lasers [28, 29], large scale spatiotemporal systems [30, 31, 32], and random projections [33] to name a few.

- Spiking neural networks: This direction of research can be seen as a specialized case of coupled nonlinear nodes with a small modification; we use a specific nonlinearity *vidilicet* excitability that is found in neurons. It is regarded that the one of the plausible causes for energy efficiency of the brain is the efficient neural coding, which this direction of research seeks to leverage to build new computing primitives. The essential element for spiking neural networks are excitable primitives. There have been many optical primitives that demonstrate excitability and thus in principle can be used to create spiking neural networks. Some experimental demonstrations of spiking behaviour are shown in are shown in refs. [34, 35, 36, 37, 38, 39].

A more complete list of bibliography on optical computing can be found in the following comprehensive reviews [40, 41, 42].

1.1 Outline of the thesis

In this thesis, we will present advances made in the field of coupled nonlinear nodes and more specifically, in spiking neural networks. The optical primitive we use is a micropillar vertical cavity surface emitting laser with an integrated saturable absorber (VCSEL-SA); this primitive will be henceforth referred to as a micropillar laser. The main results include the dynamics of micropillar laser(s) in the presence of delayed optical feedback and spatial coupling.

In the next section of the introduction, we will demonstrate a few key neuromimetic properties of the micropillar laser previously reported by the group. This would serve as an introduction for the rest of the manuscript.

In the second chapter, we will demonstrate the dynamics observed when a micropillar laser is subjected to delayed optical feedback. We will also present therein experimental results and detailed numerical analyses to describe the regimes in a qualitative and quantitative manner.

In the third chapter, we will illustrate numerically the possible integrated circuits that can be built using evanescently coupled micropillars. The last section in this chapter will be devoted to describing a computational framework using temporal coding implemented on an ensemble of micropillar lasers.

The fourth chapter presents numerical techniques based on machine learning and information theory to perform cross-prediction of the occurrence of extreme events well in advance. These techniques are applied to the case of extreme events occurring in a line laser based on the same laser stack as the micropillar laser. The observed extreme events following the onset of spatiotemporal chaos were previously reported by the group.

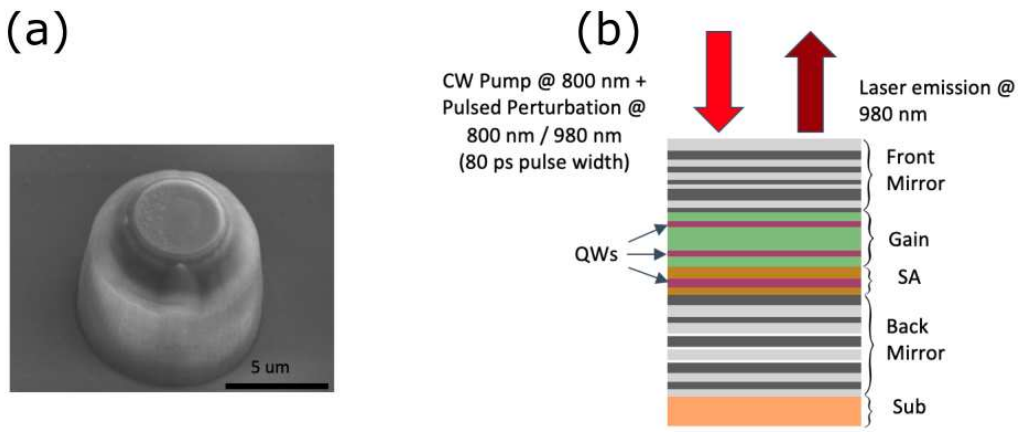


Figure 1.3: Micropillar laser
(a) Shows a scanning electron microscope (SEM) image of the micropillar laser fabricated with a SiN cladding. (b) Shows the laser stack with a few typical characteristics. (b) Adapted from [35]

1.2 Micropillar laser

Excitable micropillar laser with a saturable absorber were first demonstrated in [35] with an original laser stack design demonstrated in [43]. The cavity is designed to have an active zone comprising of two quantum wells and a passive zone comprising of one quantum well respectively. The three quantum wells (InGaAs/AlGaAs) are identical with minor differences in the barriers. The cavity is designed in such a way that the active zone is physically located at the maxima of the optical field in the pump window (790 nm - 810 nm) and the passive zone is located at the electric field minima in the pump window. However, all the three quantum wells are located at the electric field maxima at the cavity resonance wavelength. This effectively enables selective pumping in the active zone and the passive zone functions like a saturable absorber. Detailed notes on the fabrication of these devices can be found in chapters 2 and 3 of ref. [44]. The end result is a micropillar laser having a diameter of 5 μm and height of approximately 7 μm. The exact height depends on the depth of etching of the micropillar. We stop the etching leaving a few layers of the back mirror to avoid absorption of the pump into the GaAs substrate which can cause parasitic heating. Figure 1.3 shows in (a) an SEM image after the fabrication process and in (b) a schematic and typical characteristics of the micropillar laser.

Following that, several experiments supported by numerical simulations have revealed

that the micropillar laser is capable of demonstrating a variety of neuromimetic properties *videlicet* refractory period(s), temporal summation, and spike latency. The Yamada model with spontaneous emission used to simulate this system was introduced in ref. [45] and has been shown to be in excellent qualitative agreement with experimental findings. It was further shown in ref. [46] that this model is an analog to the leaky integrate-and-fire model. The model equations are as follows[2]:

$$\begin{aligned} \dot{G} &= \gamma_G(A - G - GI) \\ \dot{Q} &= \gamma_Q(B - Q - sQI) \\ \dot{I} &= (G - Q - 1)I + \beta_{sp}(G + \eta_1)^2 \end{aligned} \tag{1.1}$$

The model consists of three nonlinear coupled ordinary differential equation for the intracavity intensity (I), and the scaled excess carrier densities with respect to transparency in the gain (G) and the saturable absorber region (Q). A and B correspond to the pump intensity and the nonsaturable losses. The saturation parameter is $s = \frac{a_Q \gamma_G}{a_G \gamma_Q}$, where $a_{G,Q}$ is the differential gain and differential absorption respectively. The spontaneous emission factor and transparency offset of gain are β_{sp} and η_1 respectively. The scaled carrier recombination in the gain and saturable absorber region are $\gamma_{G,Q}$ respectively. In this slow-fast system, since the carriers dynamics are slower than the intensity dynamics, we have $\gamma_{G,Q} \ll 1$. All the timescales here are normalized to the cavity photon lifetime which is approximately 1.3 ps by the cavity design. A necessary condition for excitability is that $s > 1 + 1/B$ [47]. This is satisfied as by construction as we have $\gamma_G \approx \gamma_Q$ and $a_Q > a_G$. In the absence of the spontaneous emission term ($\beta_{sp} = 0$), the system admits $I = 0$ as an invariant manifold. Thus, any perturbation on G or Q does not increase the intracavity intensity[3]. To account for experimental findings, the model incorporates a small spontaneous emission term β_{sp} which is typically 10^{-5}. The addition of this term slightly alters the steady state values of all the variables which at order 0 in β_{sp} take the form $(A, B, 0)$ for (G, Q, I) respectively. These modifications are shown in ref. [45]. Based on

[2] We will reintroduce the model throughout the manuscript where ever it is used for the ease of the reader

[3] $\dot{I}$ is trivially 0 if $I = 0$ which prohibits any change in the intensity

this it was estimated in ref. [48] that the excitable regime (below laser threshold) corresponds to $A < 1 + B$. The neuromimetic properties arising from the micropillar laser are summarized in the subsequent subsections.

1.2.1 Excitability

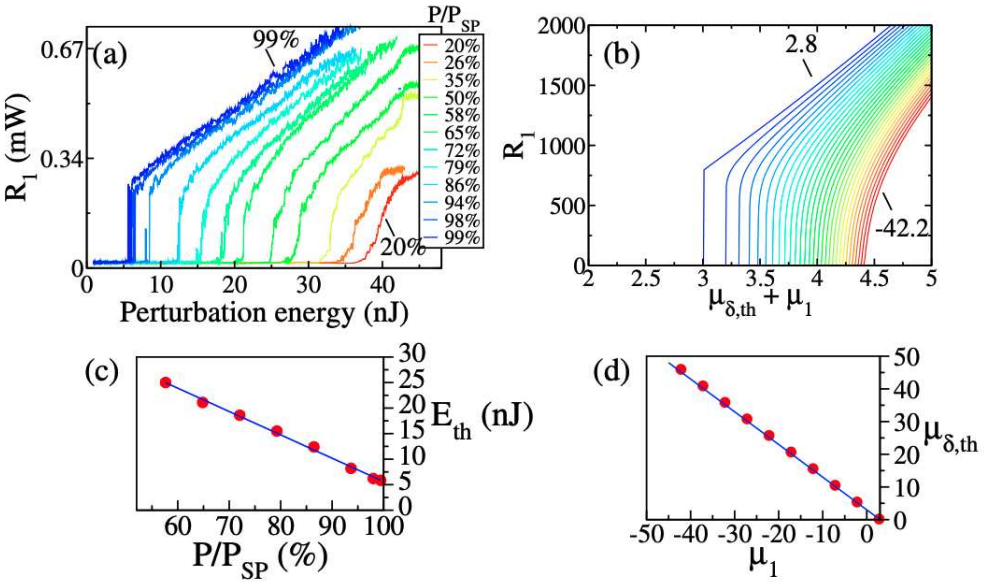


Figure 1.4: Experimental and numerical study of excitable behaviour in micropillar laser with incoherent excitation

Subplot (a) shows amplitude of the response R_1 to a single pulse perturbation versus perturbation energy E for varying bias pump P relative to the self-pulsing threshold $P_{SP} = 694$ mW. Subplot (b) shows theoretical response amplitude R_1 to single input δ-perturbation pulse μ_δ for different bias pumps $\mu_1 \equiv A$ ranging from 2.8 to – 42.2. Subplot (c) shows the dependence of the excitable threshold E_{th} (red circles) with reduced bias pump P/P_{SP} and linear fit (blue line). Subplot (d) shows excitable threshold μ_δ versus bias pump A. The blue line is the theoretical approximation given by $1 + B - A$. The parameters used for the numerical simulation were: $\gamma_G = 0.001$, $\gamma_Q = 0.002$, $B = 2$, $s = 10$, $\eta_1 = 1.6$ and $\beta_{sp} = 10^{-5}$. Adapted from [48].

It was shown that the micropillar laser demonstrates a neuron like excitability with the emission of a sub-nanosecond response [48, 49] when perturbed by a single optical perturbation on the gain or intensity variable. This excitable behaviour is characterized by a typical all-or-none response when subjected to a perturbation. The perturbation on the gain variable are carried out using a pulsed laser operating in the pump window and is called an incoherent perturbation. A perturbation carried out at the cavity resonance

is called a coherent perturbation and affects the intensity variable. The experimental and numerical results for the former are presented in figure 1.4. Below a certain threshold, called the excitable threshold, a perturbation results in no response. However, a perturbation just above this threshold generates a response. This sharp threshold is a characteristic of excitability.

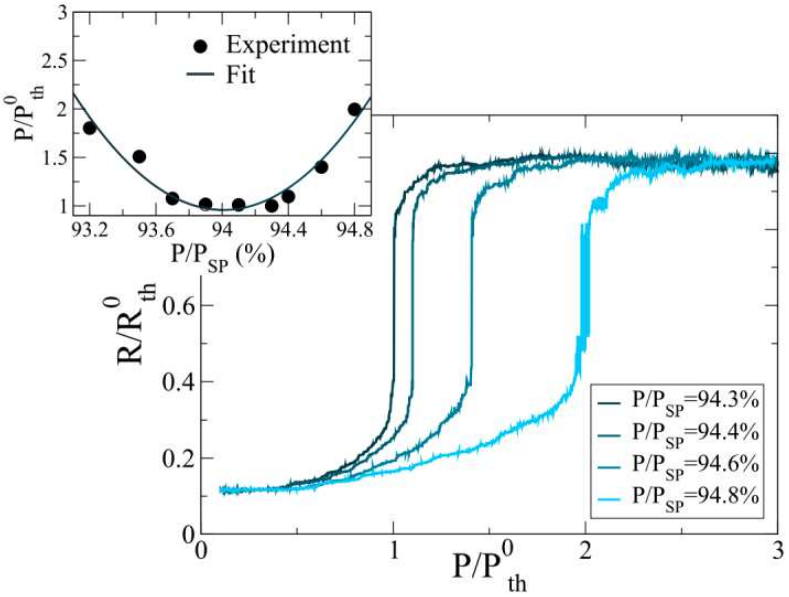


Figure 1.5: Experimental study of excitable behaviour in micropillar laser with coherent perturbation

Left panel: Response amplitude for a coherent perturbation at $\lambda = 980.47$ nm and different bias pumping with respect to the self-pulsing threshold P_{SP} pumping value. The response and perturbation amplitudes are scaled to their maximum value for $P/P_{SP} = 94.3$. Upper left panel: Excitable threshold dependence for coherent perturbations versus bias pump.. Adapted from [49].

For an incoherent perturbation, it has been demonstrated that following the sharp transition at the excitable threshold, the emission from the micropillar laser exhibits a linear growth with perturbation strength. The excitable behaviour is observed over a range of pump (A) value below the self-pulsing threshold. However, reduction of the pump value beyond a certain value causes the excitable character to disappear and to lead to just gain switching.

The response to a coherent perturbation, while having the same sharp transition, is subtly different. The results for coherent perturbations are shown in figure 1.5. In this case, following the excitable threshold, the output is clamped to a value and is no longer de-

pendent in the input. It corresponds more closely to the standard expected excitable behaviour. The two types of excitable response vs perturbation method were analyzed in detail in Ref. [50]. The different curves correspond to different values of the pump. In this case however, the excitable threshold seems to increase with the pump as opposed to the trend displayed by the incoherent perturbations. This counter-intuitive trend is due to an experimental technicality since the perturbation laser wavelength is fixed in this experiment. The changing of the pump induces thermal detuning of the cavity resonance. This in turn increases the excitable threshold. It was estimated experimentally that the excitable threshold (excluding the pump) for coherent and incoherent perturbation was approximately 4 fJ and 725 fJ respectively. Whereas was the excitable response was on the order of 50 fJ. The typical FWHM pulse duration of the excitable response was estimated to be $\lesssim$ 200 ps.

1.2.2 Refractory periods

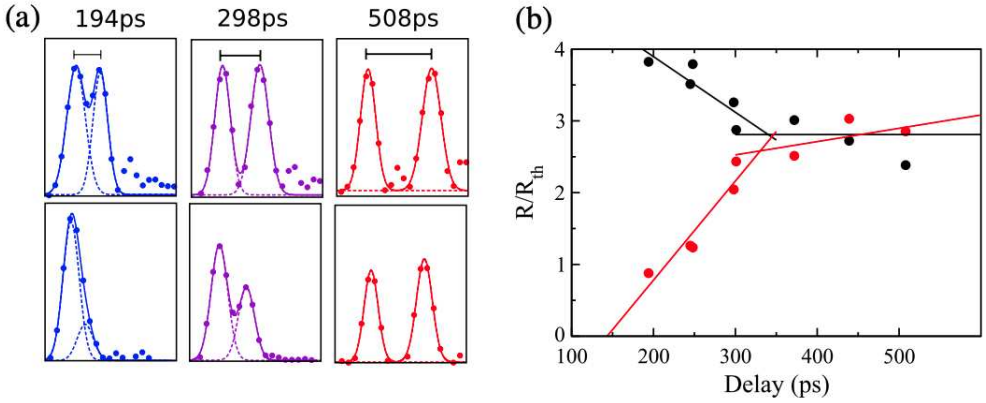


Figure 1.6: Absolute and relative refractory periods
Subplot (a) shows the recorded time traces for different delays and their Gaussian fits. Upper traces are the input perturbations and the lower traces are the system's response. The bias pump is set to 71% of the self pulsing threshold. (b) Amplitude of the response R to the first (black) and second (red or gray) perturbation pulses for a double-pulse perturbation with variable delays. R_{th} is the response amplitude at the excitable threshold. Lines are linear fits in selected ranges and are guides for the eye. Adapted from [48].

The results reported in this subsection were published in ref. [48]. In the presence of two or more perturbations, three regimes of responses were observed *videlicet* absolute re-

fractory period, relative refractory period, and an independent response. The last regime is a trivial observation of excitability when the time difference between perturbations is far greater than any internal time scale of the system. These results can be seen in figure 1.6. For this experiment, the micropillar laser was perturbed twice incoherently; each perturbation was greater than the excitable threshold so as to elicit a response by itself. The spacing between the two perturbation was between 194 ps and 508 ps. It was observed that for very short delays (< 190 ps), only one response was recorded. This regime is the absolute refractory period as the second perturbation fails to elicit a response. When the time delay is between 200 ps and 350 ps, it is observed that both the perturbations contribute to two responses but the second response is attenuated. This attenuation is due to the time taken for the carriers to recover within the gain and saturable absorber regions. Lastly, when the timing is greater than 500 ps, the two perturbations produce almost identical responses.

1.2.3 Temporal summation

The results in this subsection were published in ref. [51]. Temporal summation is the mechanism where sub-threshold perturbations arriving within a given time window can add up to elicit an excitable response. A neuron capable of temporal summation is also called a coincidence detector neuron. Coincidence detector neurons play an important role in neuromorphic computing and are thought to enable azimuthal sound localization. The experiments were carried out using coherent and incoherent perturbations. For the incoherent perturbations, two perturbations with amplitudes below the excitable threshold (74% and 80% of the excitable threshold) were sent to the micropillar laser. Individually, both the perturbations would not elicit a response from the micropillar laser. These two perturbations are sent to the micropillar with different delays and the response is recorded from the micropillar laser. The results of this are shown in figure 1.7. Since noise in the form of pump noise and spontaneous emission is present in the system, the experiment was repeated 10,000 times for statistics. For delays shown in subplot (a) - (e), the micropillar performs temporal summation with different success rates (99.97% for subplot (a) and 84.5% for subplot (e)). The temporal profile of the excitable response in each

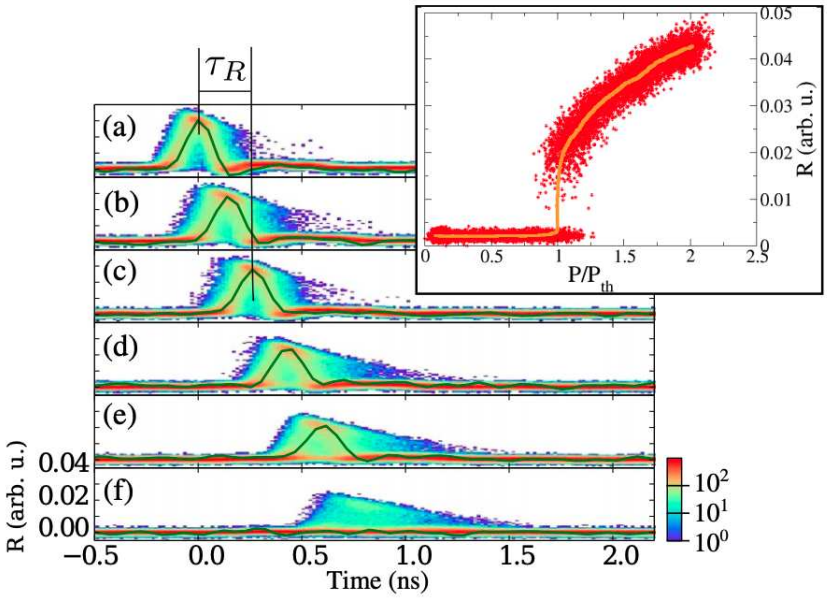


Figure 1.7: Temporal summation with incoherent perturbations
Experimental traces of the system's response to two incoming, sub-threshold perturbations for different perturbation delays δ: (a) 210, (b) 320, (c) 420, (d) 520, (e) 610, and (f) 700 ps. The plots show the statistical density of points (in log scale) for 10,000 different realizations. On the plots a typical response pulse is shown in green. Inset: excitable response to a single perturbation. Red stars are the detected response maxima. Orange is a plot of the median in a sliding window with 500 points. Adapted from [51].

case is almost identical. However, in subplot (f) when the delay is 700 ps, the micropillar laser fails to perform the temporal summation with a high success rate.

For the coherent perturbation experiment, the two perturbation were fixed to be 44% and 66% of the excitable threshold. In this case the temporal summation happens for an shorter temporal window. The results from this experiment are shown in figure 1.8. In subplot (a) and (b), where the delay was 220 ps and 350 ps respectively, the temporal summation occurs; the success rate in (b) is lower. For other cases subplots (c)-(e), the temporal summation does not occur. The locations of both the perturbations are marked by arrows in the figure.

In both the coherent and incoherent temporal summation, there is a delay which scales nonlinearly with respect to the individual perturbations and the spacing between them. This nonlinear delay could be an interesting mechanism for implementing temporal codes. Similar mechanisms for temporal summation can be implemented using multiple evanes-

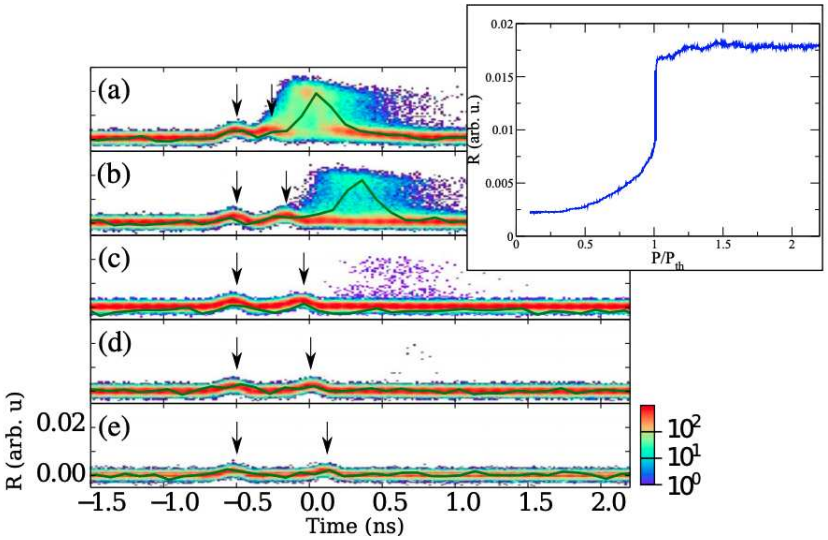


Figure 1.8: Temporal summation with coherent perturbations
Experimental traces of the system's response to two incoming, sub-threshold perturbations for different perturbation delays δ: (a) 220, (b) 350, (c) 450, (d) 540, and (e) 630 ps. The plots show the statistical density of points (in log scale) for 10,000 different realizations. On the plots, a typical response pulse is shown in green. The intensity perturbations are indicated by arrows. Inset: excitable response to a single perturbation median-averaged over 500 points. Adapted from [51].

cently coupled micropillars. These will be presented in chapter 3.

1.2.4 Spike latency

Spike latency is the delay between the arrival of a perturbation and the emission of an excitable response. In the micropillar laser, it is demonstrated in ref. [49] that there is a sub-nanosecond nonlinear delay that depends on the pump as well as the perturbation strength; an analytical expression was also derived. This experimentally measured nonlinear dependence for an incoherent perturbation is shown in figure 1.9. In the figure, the different curves correspond to the different values of the pump. For each value of the pump, several perturbations of different strengths were sent to the micropillar laser. For a pump value giving rise to a low excitable threshold, the measurable latency can change over a large range (150 ps to 1 ns). This inverse relation between the perturbation strength and response time can enable the use of micropillar lasers for temporal coding which is the coding of information through the timings between spikes. More results will presented on this will be presented in chapter 3.

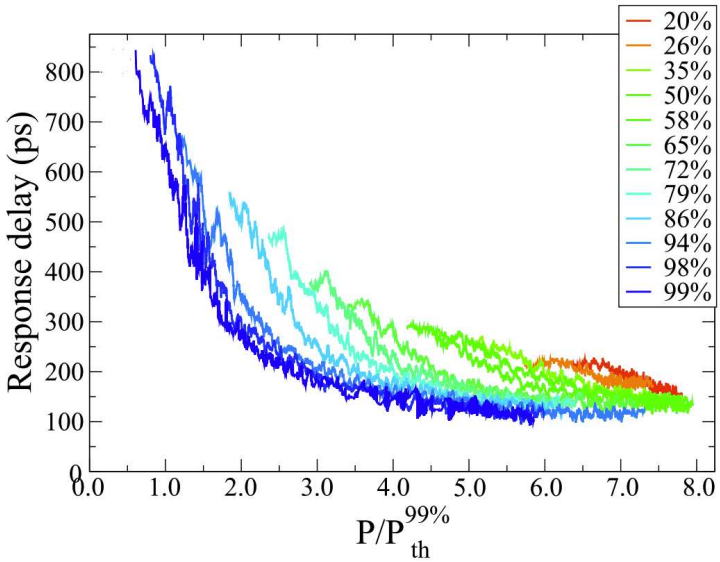


Figure 1.9: Spike latency of a micropillar laser with incoherent perturbations
Median of the pulse response delay versus normalized perturbation energy P for different bias pumps with respect to the self-pulsing pump threshold. The perturbation and the response are normalized respectively to the excitable threshold ($P_{th}^{99\%}$) and to the response at excitable threshold $R_{th}^{99\%}$ for a bias pump P equal to 99% of the pump at the self-pulsing threshold. Adapted from [49].

Chapter 2

Delay Coupling

Having described the context of our research and introducing the laser system under investigation in chapter 1, this chapter is dedicated to delay-based architectures from a fundamental as well as an applied point of view.

Delay coupling refers to a time delayed feedback to a single micropillar laser by the means of an external cavity as shown in figure 2.1. We first present the experimental setup and introduce the model equations used. Following this, we show the existence of regenerative pulse trains in the system and demonstrate how we can manipulate them with single optical pulses. This demonstrates the capability of optical buffer in the short term. We proceed to study the asymptotic dynamics of the system which consists of periodic and symmetry-broken states. All the experimental results are supported by numerical and theoretical analysis in collaboration with our partners. We finish by discussing the implication of these results in the context of neuromimetic information processing.

2.1 Experimental Setup

The micropillar laser is optically pumped to be in the excitable regime with a continuous wave laser (Coherent FAP system) in the pump window (between 790 nm and 810 nm) and emits light at the cavity resonance wavelength (between 960 nm and 990 nm). The CW pump is controlled by an DC voltage source. The pump profile on the sample plane

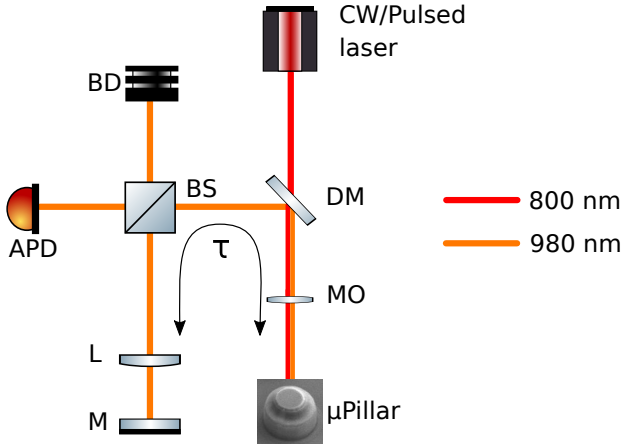


Figure 2.1: Schematic of the experimental setup with time delayed feedback
Figure shows the optically pumped excitable micropillar laser with delayed optical feedback from an external mirror. DM: Dichroic mirror, BS: Beam splitter with 70/30 power split between reflected and transmitted path, MO: Microscope objective, APD: Avalanche photodiode, L: Lens with f = 5 cm, M: High reflectivity feedback mirror, BD: Beam dump, μPillar: Micropillar laser, τ : External cavity round trip time.

is flat top and power density[1] at threshold is on the order of 10 - 100 $\mathrm{kW/cm^2}$; the exact value depends on individual sample properties (position of the cavity resonance with respect to saturable absorber bandgap) and also varies in different fabrication runs. The micropillar laser is cooled and stabilized to a temperature slightly below 0°C using a Peltier element with a stability better than 0.1°C. The output light from the micropillar is split using a R/T = 70/30 beam splitter (BS). The transmitted part is detected using a 5 GHz bandwidth avalanche photodiode (APD), amplified by a large bandwidth (18 GHz) RF amplifier and analysed with a 13 GHz oscilloscope. The reflected part is directed into an external cavity, closed by a high reflectivity mirror (M) after focusing with a 5 cm focal length lens, which provides a tunable delay τ on the order of 10 ns. This delayed optical feedback results in a 10% reduction of the laser threshold. The micropillar laser is perturbed by short optical perturbations of 80 ps duration from a mode-locked Ti:Sa laser (Spectra Physics Tsunami). The repetition rate of the Ti:Sa laser is fixed by the cavity round trip time and is about 82 MHz (12.28 ns period). This is thus the minimum time duration between two perturbations. At the laser output there is a pulse picker which can select pulses such that the repetition rate can be lowered. The perturbations can either

[1]The typical diameter of the pump beam using an 80x microscope objective is 10 μm

be coherent or incoherent with respect to the cavity resonance wavelength. Both classes of perturbations result in similar responses with subtle differences.

2.2 Model Equations

We model our system using the Yamada model as shown in the previous chapter (Eq. 1.1) with $\beta_{sp} = 0$ and adding time delayed feedback term. The equations are as follows:

$$\begin{aligned} \dot{G} &= \gamma_G(A - G - GI) \\ \dot{Q} &= \gamma_Q(B - Q - sQI) \\ \dot{I} &= (G - Q - 1)I + \kappa I(t - \tau) \end{aligned} \tag{2.1}$$

κ and τ are the feedback strength and delay, respectively. In a strict sense, this model is incapable of responding to incoherent perturbations (perturbations on G) as at steady state it admits $I = 0$ as an invariant manifold. A spontaneous emission term could be added but we take $\beta_{sp} = 0$ to ease the theoretical analysis. Thus for incoherent perturbations to act effectively, we just have to take a small non-zero initial value for intensity I.

This model considers the intracavity laser intensity I instead of the electric field E essentially discarding phase effects for the following reasons:

- Our system is in the excitable regime with the stable state as the laser off state. Therefore, in this off state, the only optical field present is due to spontaneous emission which is not coherent and is of very low intensity.
- The delay time τ is typically one order of magnitude greater than the FWHM of the emitted pulse ($\lesssim$ 200 ps) which corresponds to the time duration of the excursion in the phase space. This then ensures that the feedback always arrives when the laser is in the off state.
- The model gives good qualitative agreement with the experimental observations. A comparison of the Yamada model with a model including phase effects [52] yielded

effectively similar results in the regime considered by us.

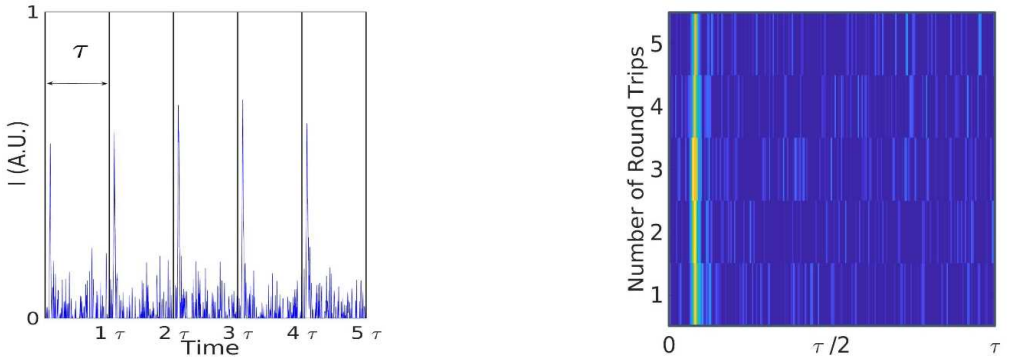
2.3 Manipulation of pulse trains in delay systems

We set the micropillar laser in the excitable regime: the steady state intensity I is zero, but a single high-amplitude, short pulse of light can be emitted in response to an external perturbation of sufficient amplitude [35, 53]. When subject to delayed feedback, an excitable system can either

- Remain in its off state for external perturbations below the excitability threshold.
- Emit a single pulse if the perturbation exceeds the excitable threshold but the external cavity losses are too high.
- Regenerate its own initial excitable response after the reinjection time τ giving rise to periodic pulse train for a sufficiently high feedback strength. The repetition rate of the regenerative pulse train is close to τ[2].

A useful tool to analyse such pulse trains is a two dimensional pseudospace representation [55, 56]. We would like to note that the pseudo-space representation we use here does not necessarily mean that the vertical (y) axis can be treated as an extra spatial dimension. In fact, for this to be true we need very long delays compared to the correlation time in the system. We just use it because it is convenient to follow the evolution of pulses. This 2-D representation facilitates visualizing the information easily for several pulses per round trip over many round trips. The generation of such two dimensional representations from time series follows a simple folding operation: the time series data (figure 2.2a) is divided into segments of τ and these individual segments form the rows of a matrix as depicted in figure 2.2b. The x-axis of figure 2.2b is now *continuous time* between 0 and τ and the y-axis is the number of round-trips or *discrete time*.

[2]It is close to but not equal to τ as the emission period is the sum of the delay and spike latency. In the previous works from our group, we showed that such a system can sustain optical temporal dissipative solitons for a finite duration. This finite duration results from a pump noise-induced escape from a stable periodic solution to a stable equilibrium [54]



(a) Time trace (b) 2D Pseudospace representation

Figure 2.2: Transformation from time traces to 2D pseudospace representation

This very general mechanism for self-pulsations[3] has been implemented in different optical systems, including a coherently driven vertical-cavity surface-emitting laser (VCSEL) [37], a VCSEL subject to optoelectronic feedback [57], two coupled semiconductor lasers [58], a photonic resonator with optoelectronic self-feedback [59], and a micropillar laser with integrated saturable absorber [54]. Since almost arbitrary pulse timing patterns can, in principle, be excited and regenerated after each delay, regenerative dynamics can be of particular interest for producing complex optically controllable temporal pulsing patterns [60, 61, 62, 63, 64] or for spike-based optical memory applications [37, 59, 64, 65].

It has been shown that the pulsing dynamics is solely governed by the nonlinear spike latency –time between reinjection and emission of a pulse-– which is a function of the net gain $\widetilde{\mathrm{G}} = G - Q - 1$ accounting for the gain and saturable loss at the reinjection time. In the vicinity of the perturbations, we report similar behaviour as demonstrated by past works [37, 59] in showing an all-optical information storage system which can perform noise correction[4] by the virtue of its excitable behaviour. We further present an all-optical control over the information in the buffer either using the CW pump or a single optical perturbation pulse.

The all-optical control can be presented as addition of a pulse, retiming of pulse train, tweezing of a pulse train in figure 2.3. In all the subplots here, the red circle indicates

[3]This mention of self-pulsations is different from the self-pulsing threshold of a single micropillar laser. The latter corresponds to the homoclinic bifurcation in a single micropillar laser without feedback and is a result of the interplay between the variables in the Yamada model. 2.1

[4]Here noise correction refers to the fact that on every feedback instance, the output is a fixed response as long as the input exceeds a threshold thus avoiding any accumulation of noise.

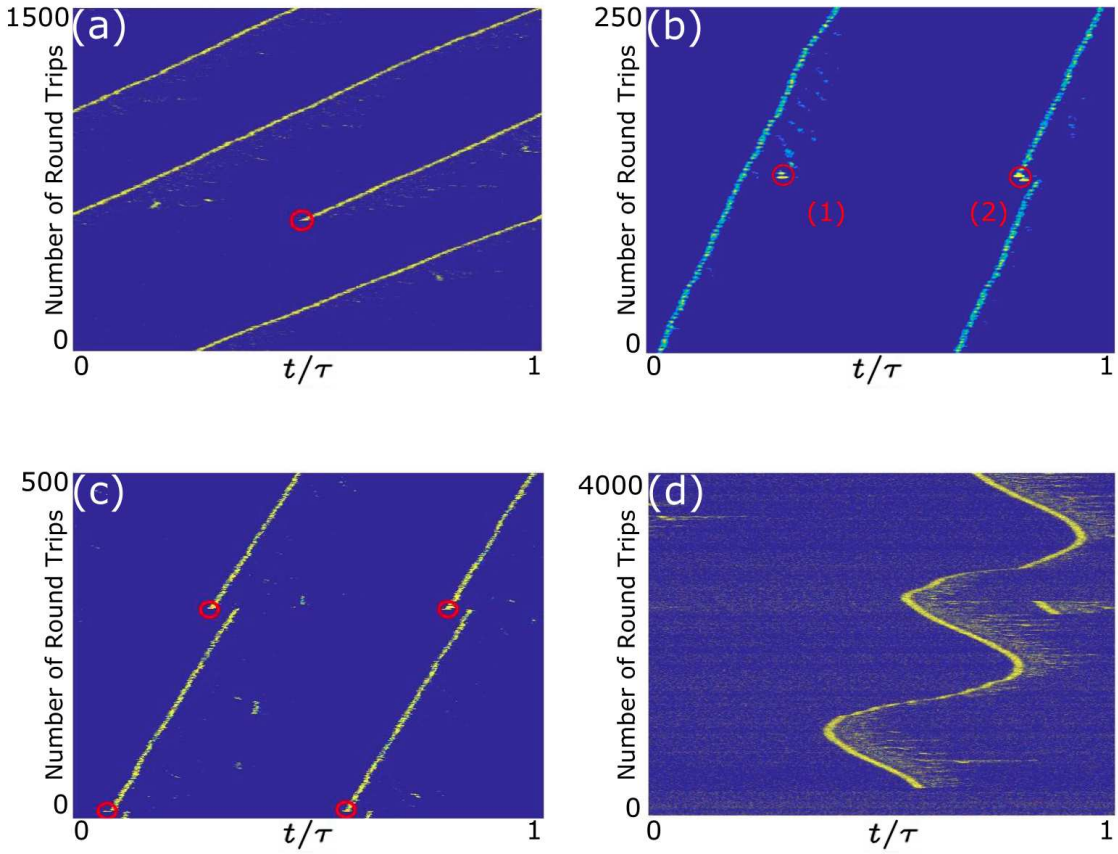


Figure 2.3: Operations on the all optical buffer
In all the panels here, the red circle indicates a perturbation. In subplot (a) we see the addition of a pulse train, in subplot (b) we see the effect of refractory period in perturbation (1) and retiming of a pulse train due to perturbation (2), in subplot (c) we see the retiming of a pulse train and in subplot (d) we see the tweezing of a pulse train. See text for additional details.

a single optical perturbation. Further, we represent the 2-D space with the two coordinates with the convention that the first coordinate corresponds to $\frac{t}{\tau}$ and the second coordinate corresponds to the roundtrip number. In order to manipulate the pulse train we need to have a proper control over the perturbation times. This is achieved by the use of an external trigger to the pulse picker of the Ti:Sa mode-locked laser. The external trigger (generated using Agilent 8114A pulse generator) is a window of several tens of nanoseconds which controls the acousto-optic modulator of the pulse picker. This way, we can choose a given sequence of perturbation pulses from the train of pulses of the Ti:Sa mode-locked laser. In order to achieve a proper timing of the perturbation pulse with respect to the existing pulse train, we need to account for the two time scales: delay

time τ and the repetition time of the Ti:Sa mode-locked laser (12.28 ns). On a technical note, the pulse from the external trigger is larger than the mode-locked repetition period. Thus, to select a single pulse, we need to properly adjust the time of external trigger (accounting for the rise and fall time of the trigger pulse) with respect to the pulse train.

In subplot figure 2.3(a) a pulse train with one pulse per round trip is first excited. We then send a second perturbation which is timed such that it is not in the vicinity of the existing pulse train. We observe the addition of a pulse train when we perturb the system at point $(\mathrm{t}_2, \mathrm{t}_1)$ with an existing pulse train arriving at $(\mathrm{t}_3, \mathrm{t}_1)$ such that $\mathrm{t}_2 < \mathrm{t}_3$ and $|\mathrm{t}_3 - \mathrm{t}_2| > \mathrm{t}_{\mathrm{refractory}}$ [5].

In subplot (b) we discuss the effect of perturbation labeled as (1). The perturbation (2) causes a retiming of a pulse train and this mechanism is discussed in the next figure. Perturbation (1) is sent at point $(\mathrm{t}_2, \mathrm{t}_1)$ with an existing pulse train at point $(\mathrm{t}_3, \mathrm{t}_1)$ such that $\mathrm{t}_3 < \mathrm{t}_2$ and $|\mathrm{t}_3 - \mathrm{t}_2| < \mathrm{t}_{\mathrm{refractory}}$. Due to this, the perturbation fails to start a second pulse train and has no effect on the system. In subplot (c) we follow a procedure similar to (a) but in this case, the difference $|\mathrm{t}_3 - \mathrm{t}_2|$ is less than $\mathrm{t}_{\mathrm{refractory}}$ resulting in the existing pulse train having insufficient gain to regenerate effectively retiming the pulse train. The slope of the mean trajectory observed in all the panels is a result of the imprecise estimate of the folding time. The folding time is taken in this thesis, for practical reasons, as the fundamental pulse repetition period which is estimated using the fundamental peak in the Fourier spectrum of time traces. The folding time is set to be as close to this time as possible. Since the resolution in the folding time is set by the sampling time of the oscilloscope, we cannot have arbitrary precision. This lack of precision manifests as slight drifts in the pseudo-space representation and gives rise to the slope.

In subplot 2.3(d) we perturb the system only once resulting in a pulse train, we then modulate the pump (CW) laser using a sinusoidal signal generated by Agilent 33220A arbitrary waveform generator which in turn drives the current in the FAP pump laser system. The sinusoidal signal has a mean of 530 mV and a modulation of 70 mV V_{pp} with a frequency of 80 kHz. The resulting change in net gain is manifested in the form of changing spike latency giving the effect of tweezing.

[5] $\mathrm{t}_{\mathrm{refractory}}$ is the refractory period.

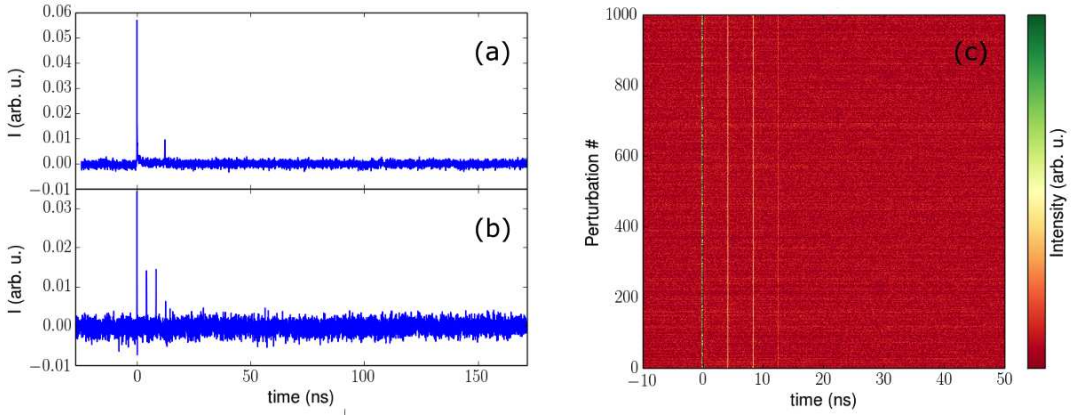


Figure 2.4: All optical control of a pulse train
In subplot (a) we see the perturbation sequence, in subplot (b) we see the response and in subplot (c) we see the same experiment repeated thousand times.

Using similar mechanisms it was shown in [64] that it is also possible to switch on and off pulse trains in the system with two single (incoherent) optical perturbations. In figure 2.4, subplot (a) shows the perturbation sequence and subplot (b) shows the response recorded from the system. It was demonstrated in chapter 1 fig. 1.4 that in the case of incoherent perturbations, we observe a sharp threshold in the response at the excitable threshold and then a linear growth in response with respect to the perturbation strength. The first large perturbation in (a) triggers the pulse train shown in (b) but the subsequent perturbation generates an insufficient response in the system while arriving just before the regeneration of the fourth pulse in the pulse train. This second perturbation uses up the gain in the system and the existing pulse train falls in the refractory period, switching it off. The response from the second pulse perturbation is not high enough to reach the excitable threshold after suffering cavity roundtrip losses and thus it generates no pulse train. This interaction makes the switch on and off of pulse trains using single optical perturbations. Subplot (c) shows the results of this experiment carried out thousand times with the $\tau \approx 4.2$ ns; the case number is on the ordinate and time is represented on the abscissa with $t = 0$ being the first perturbation. Based on this subplot we conclude that the effect is very reproducible and not prone to experimental noise. Moreover, it can be noted that the jitter in the regenerated pulses is low (shown by the vertical lines) as it is fixed by the delay in the system.

Similar manipulation of optical pulses has been reported in ref. [62] in a passively mode-

locked semiconductor laser using electrically addressed short perturbations on the pump. They reported the switch on and off of single pulses with repeated (more than hundred) and precisely timed perturbations. In our case, we achieve pulse manipulation using single optical perturbation.

2.4 Asymptotic dynamics

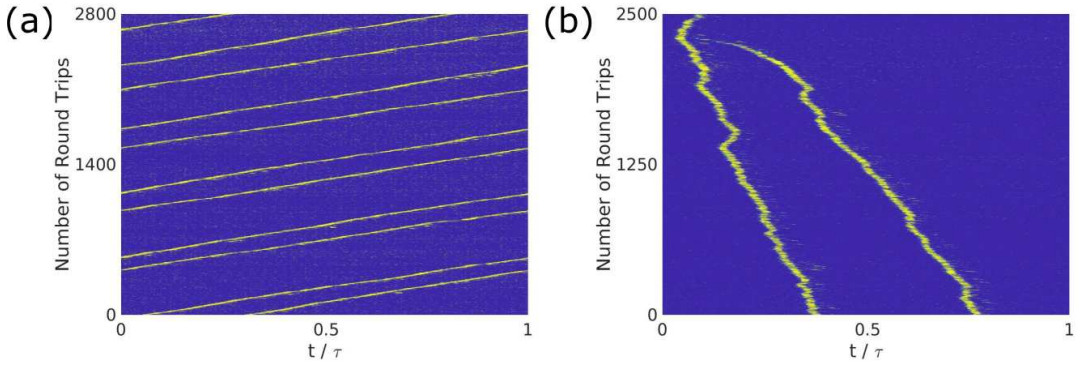


Figure 2.5: Pulse - Pulse interaction over several thousand roundtrips
In subplot (a) the feedback time is $\tau = 8.2$ns and in subplot (b) the feedback time is $\tau = 6.33$ns

While in the short term, we observe that the system behaves like an optical buffer, we observe that the pulses do interact with time and this requires further analysis. In figure 2.5, subplot (a) two initially close pulses in the external cavity tend to separate over thousands of round trips and experience a repulsive interaction. By contrast, in subplot (b), two initially separated pulses seem to experience an attractive interaction until one pulse train is turned off as it enters the absolute refractory period of the first pulse train. The two plots were obtained using different micropillars operating in different experimental conditions. To the best of our knowledge, pulse attraction in such systems was not reported before. In the next section, we further our analysis by comparing the experimental results to numerical simulations with appropriate parameters and distinguish the results based on the carrier recombination rates in the gain and saturable absorber regions. The theoretical results presented in this chapter have been developed in collaboration with Soizic Terrien, Neil G. R. Broderick and Bernd Krauskopf at the Dodd-Walls Center for Photonic and Quantum Technologies, The University of Auckland, New Zealand.

In the context of biological spiking neurons, delayed self-connections have also been recognized to play a central role in the persistent regeneration of input stimuli [66, 67, 68]. While in the short term, the regeneration occurs with minimal loss of timing information, we show that in the long term this is not true anymore. The intuitive picture of regenerating arbitrary pulse trains while accurate in predicting the short-term behaviour of the system is overly simplified. A theoretical analysis predicts more complicated dynamics, including the co-existence of several self-pulsing modes with the stable off-state [64, 69, 70]. It is well known that delay differential equations (DDEs) have an infinite dimensional phase space and can display rich dynamics with coexistence between different types of attractors [71, 72]. Thus, it is a fundamental question to understand the long-term dynamics of such systems.

2.4.1 Time symmetric pulsing patterns - Faster saturable absorber

In order to study these long term dynamics, we revert to the simulations using the equations introduced in Eqs. 2.1. The first set of parameters we consider are $A = 2.4$, $B = 2.2$, $\gamma_G = 0.01$, $\gamma_Q = 0.02$, $s = 5$, $\kappa = 0.05$, and $\tau = 1100$. These are chosen to match the known physical parameters and the experimental observations [73, 47]: recombination timescales (γ_G, γ_Q) are on the order of few hundreds of picoseconds and the feedback time (τ) is between 5 and 10 ns. The saturable absorber recovers twice as fast as the gain. The small value of γ_G and γ_Q represents the slow timescale of nonradiative recombination in the quantum wells as opposed to the fast timescale of the cavity photon lifetime. After emitting a pulse, the net gain $\widetilde{\mathrm{G}} = G - Q - 1$ recovers to 95% of its steady state value in $\mathrm{t_{rec}} = 393$; the feedback time is thus approximately 3 $\mathrm{t_{rec}}$ which is comparable to the experiment. The model equations 2.1 do not account for pump or spontaneous emission noise, thus the focus is solely on deterministic dynamics. As described in the previous section, in the vicinity of the perturbation, the system acts like an optical buffer and preserves information in the form of inter-spike distances, here in figure 2.6 we compare two examples of experimental time traces with numerical simulations which appear to be in good agreement. In panel (a1), two successive perturbations are sent with a time difference of approximately 12.28 ns, which results in two pulse trains with an inter-

spike distance of approximately $0.4 \times \tau$. This distance is preserved over 250 roundtrips. The evolution of three coexisting pulse trains can be seen in the next panel. Here the non-equidistant pulsing pattern is shown to exist over a short duration of 30 roundtrips.

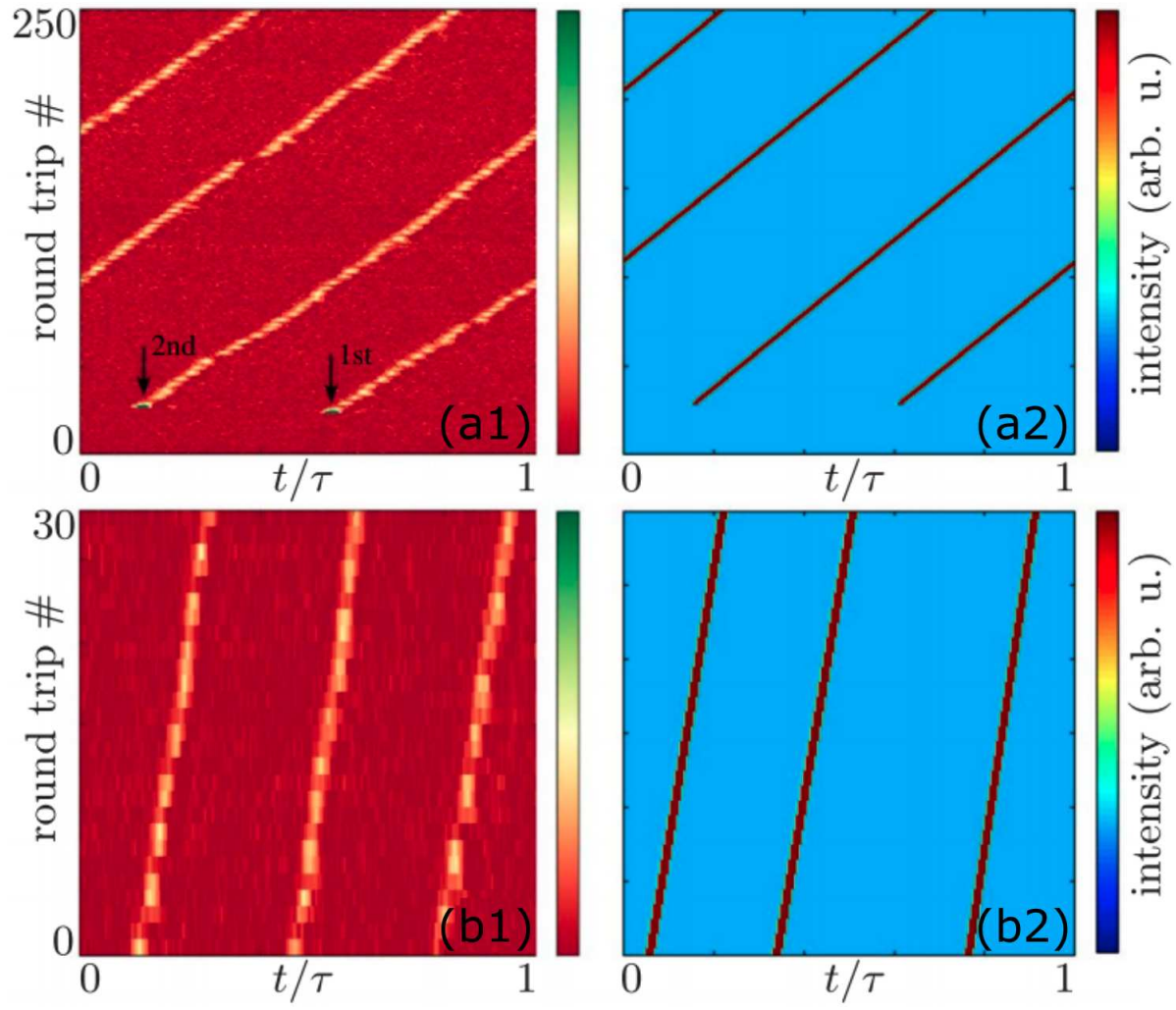


Figure 2.6: Pseudospace representation of pulse trains over short duration
Experimental (left) and simulation (right) results of pulse trains over short duration. (a) and (b) Represent two and three pulses per round trip respectively. The feedback delay was 4.77 ns in the experiment. Figure adapted from [64]

For the same parameters, figure 2.7 represents the phase portrait of system (2.1) in the (G, I)-plane, calculated with the continuation toolbox DDE-Biftool [74, 75]. It shows the coexistence of seven stable solutions: a non-lasing equilibrium and six periodic pulsing solutions. The smallest orbit corresponds to the six pulse per roundtrip solution. These coexist with several unstable periodic solutions and equilibria, which are not represented here. The stable pulsing solutions have periods T close to submultiples of τ and, hence,

correspond to different numbers of equidistant pulses in the external cavity [70]. Apart from the one for which T is close to τ, they all are only weakly stable. Importantly, there exists no stable periodic solutions with non-equidistant pulses in the external cavity, despite the fact that such solutions are observed over long periods of time in experiments and simulations. Thus the results presented in figure 2.6 only present a partial picture. They represent a transient towards one of the (weakly) attracting periodic solutions of figure 2.7. The results from the numerical integration of Eq. 2.1 with two coexisting and non-equidistant pulses are shown in figure 2.8. The convergence to the attractor only happens over several hundreds of roundtrips. The bifurcation analysis of the model shows that the amplitudes of the periodic solutions with the largest periods are very close to each other (see the two largest orbits in 2.7): as such, no significant difference is observed in the amplitudes when one or several pulses exist in the external cavity. The slow convergence towards the stable periodic solution is explained entirely by the dynamics of the net gain, $\widetilde{\mathrm{G}}$.

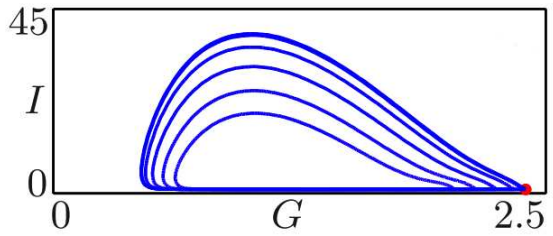


Figure 2.7: Phase portrait of Yamada model with a time delayed feedback in the (G,I) plane
Illustration of one stable equilibrium (dot) and six stable periodic solutions (curves). Figure adapted from [64]

Panels (d) and (e) in figure 2.8 plot the G and $\widetilde{\mathrm{G}}$ dynamics during 7 roundtrips (10-17) indicated by the shaded area in panel (a), the absorption recovers faster than the gain, thus immediately after a pulse, the low net gain $\widetilde{\mathrm{G}}$ increases back to its saturated value as both G and Q recover to their respective steady state values. Since $\widetilde{\mathrm{G}} = G^{\check{}}Q^{\check{}}1$ and Q recovers faster, the value of $\widetilde{\mathrm{G}}$ approaches its steady state value from below. A second perturbation is introduced in the system when the $\widetilde{\mathrm{G}}$ hasn't sufficiently recovered but the gain and perturbation strength make it possible to trigger a response. However, since this perturbation experiences a slightly lower $\widetilde{\mathrm{G}}$ it has a higher pulse latency time [49]

and hence, repetition period, for the second pulse train compared to the first pulse train. For every subsequent roundtrip this process repeats and the second pulse is reinjected further away from the first one, until both pulse trains experience an identical net gain and their repetition periods become equal. This effect of convergence due to the gain dependent latency gives rise to the apparent repulsion between the pulses. Pulse to pulse interaction is thus mediated by the carriers and is not an effect of the optical tails of the pulses. Since the system converges in the long term to a stable periodic orbit of the phase portrait the maximum number of pulse trains sustained by the system is related to the number of stable periodic solutions. Figures 2.9 show experimental results on the convergence of irregularly spaced pulse trains (two and three pulses per roundtrip) to equidistant pulse trains over the course of several thousands of roundtrips. Figures 2.9 a1, a2, b1, and b2 show the evolution of two and three pulse trains respectively using the pseudospace representation. The pseudospace plots are plotted over 200 roundtrips and at different stages in the convergence as shown in the figure 2.9 a3 and b3 which highlights the slow convergence by plotting the Δp or pulse-to-pulse timing over consecutive roundtrips. The pulse-to-pulse timing Δp slowly converges to a value close to a half or a third of the delay time τ, respectively as equidistant pulsing is approached. This slow convergence rate is on the order of a few picoseconds per roundtrip which is very small when compared to the pulse duration of approximately 200 ps. It can be observed in the experiment only over long time periods. The random-walk like fluctuations of the pulse-to-pulse timing are explained by the presence of pump noise in the system, which induces stochastic fluctuations of the micropillar net gain [54].

Having described the process of convergence of arbitrary pulse trains to equidistant pulse trains, the natural continuation would be to understand the scenarios under which it is possible to switch between these (weakly) stable solutions. From a mathematical point of view, these scenarios can be described by the basins of attraction of the different stable periodic solutions. The time traces of the six periodic solutions are shown in figure 2.10a and their periods are close to sub-multiples of the delay time, τ. For the sake of brevity, these solutions will be referred to as one-pulse solution, two-pulse solution and so on. The final state of a multistable system depends on its initial conditions. For each of the

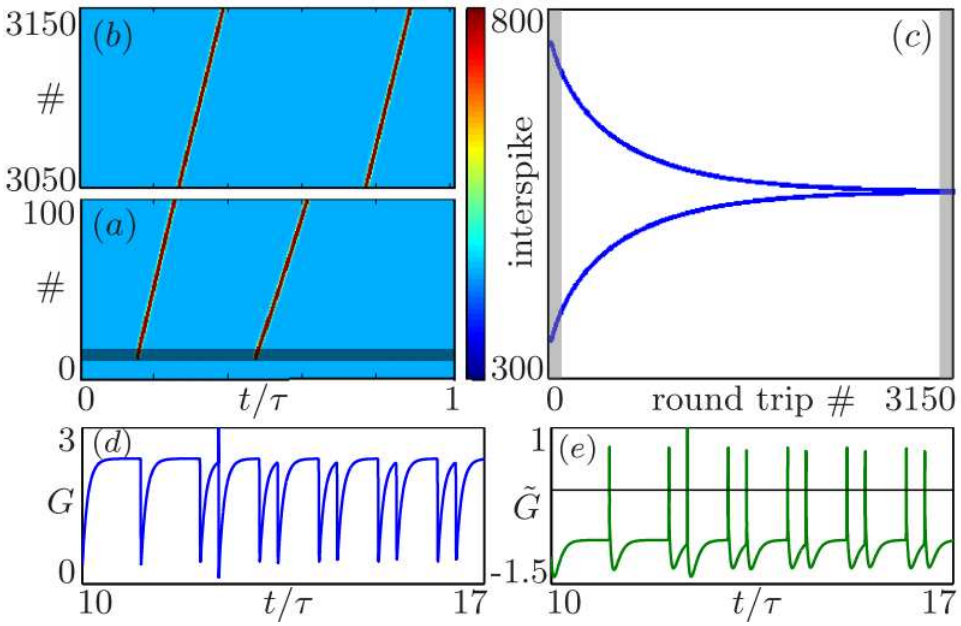


Figure 2.8: Simulation of two coexisting pulse trains
(a) and (b) Show the pseudospace representation just after the perturbations and in the long-term respectively. (c) shows the evolution of the elapsed time between successive pulses; the shaded area is the segment represented in (a). (d-e) Show the temporal evolution of G and $\widetilde{\mathrm{G}}$ respectively.
Figure adapted from [64]

stable solutions (attractors), displayed in figure 2.10a, there exists a basin of attraction. This basin of attraction corresponds to the set of initial conditions for which the transient converges to the said attractor. Thus, the basins of attraction provides us the information of how to approach different stable solutions of the system and how to switch between solutions using external perturbations. The system described here is modeled by a delay differential equation which intrinsically has an infinite dimensional phase space [76] which makes numerical continuation very complex [77, 78]. Thus we choose to integrate[6] Eq. 2.1 numerically to obtain the basins of attraction. Figures 2.10(b-g) summarize the effect of perturbations on the long-term dynamics of the system; it assumes that the system is currently in one of the stable solutions and not in a transient. A perturbation on the gain variable of amplitude ΔG at time $\widetilde{\mathrm{t}}$ relative to the existing pulse train(s) which are marked by vertical gray lines in the figure ($\widetilde{\mathrm{t}}$ = 0 is the reinjection time of a pre-existing pulse in the micropillar laser) will excite a transient but a change in the long-term dynamics is

[6]We use a custom routine based on MATLAB® DDE23 function

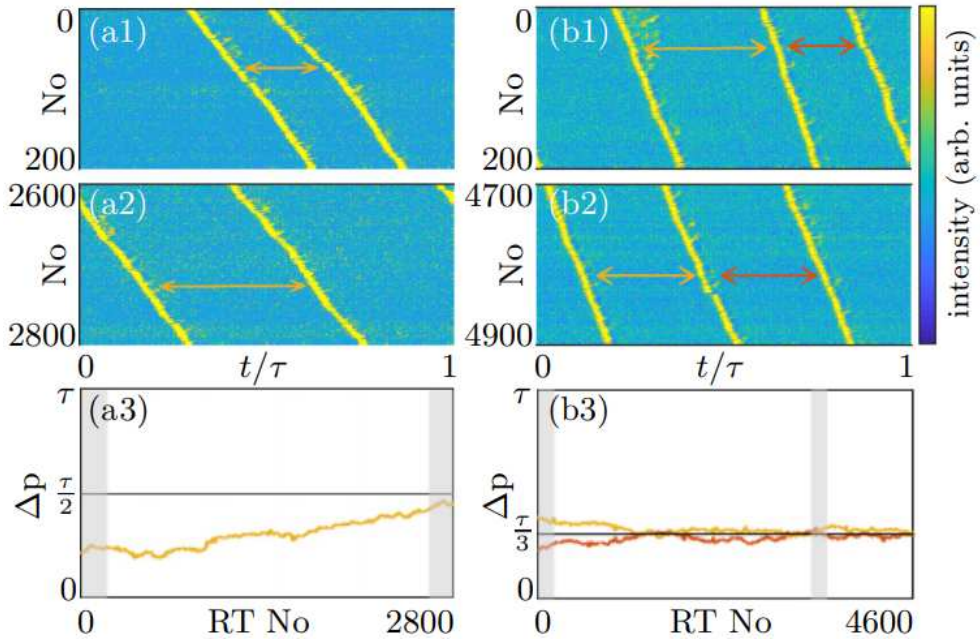


Figure 2.9: Experimental results of two and three coexisting pulse trains over a long duration

Subplots (a) and (b) Represent two and three coexisting pulse trains in the external cavity. (a1), (b1) ,(a2), and (b2) correspond to the pseudospace representation of time traces immediately after the perturbation and after several thousand roundtrips. (a3) and (b3) Plot the pulse-to-pulse timing Δp versus the roundtrip number. The shaded regions in these sub-panels correspond to the roundtrip numbers visualized in the pseudospace representations in the previous plots. The feedback delay was 8.2 ns.

conditional on perturbation being in a conducive region in the (ΔG, $\widetilde{\mathrm{t}}$) space. The grey rectangular blocks in the figures 2.10(b-g) correspond to the reinjection time of the pre-existing pulse trains. When the system is in the n-pulse regime with n $\leq$ 3 figures 2.10(b-d), a perturbation can either switch the system to the n+1 - pulse regime or can result in no change. If the perturbation is too weak (low ΔG) it might fail to regenerate in the cavity or even elicit a response from the system. However given a suitable perturbation strength the system responds depending on the timing of the perturbation; if the perturbation is in vicinity (immediately before or after) of an existing pulse train it can either cause the retiming of the pulse train or it can fall in the refractory period of the existing pulse train respectively. The effect of the relative refractory period is clearly visible in the initial negative slopes of the bottom left boundaries of the new stable pulsing regimes. Another observation from the figures is that ΔG_{min}, the minimum perturbation strength required to trigger a response increases with the number n while the time window to switch to the

n+1 pulse regime reduces.

When the initial stable regime is the n-pulse regime with $n > 3$, figures 2.10(e-g) show that any perturbation causes the system to either be in the same regime or switch to a n-1 pulse regime. As in the case of n $\leq$ 3 perturbations in the immediate vicinity of pre-existing pulses do not alter the long term dynamics of the system. Additional perturbations with appropriate timing and amplitude alter the long term dynamics of the system by removing pulses from the existing pulse train. The area of conducive parameters for removing pulses in the (ΔG, $\widetilde{\mathrm{t}}$) space increases with n. Despite these constraints, it is possible to access all the pulsing regimes in the system. The $n > 3$ pulses solution can be accessed by perturbing the system while it is still in a transient as these basins of attraction hold true only when the system starts from a stable solution. Thus if the system is perturbed n times ($3 < n \leq 6$) before it reaches a steady state, it is possible to access the n pulse solution. Numerical simulations of accessing five pulse solution are presented in figure 2.11, the two scenarios in juxtaposition are: (a) perturbations sent to a system in a four pulse solution and (b) perturbations sent to a system in a transient. In panel (a) both the perturbations fail to effect any change in the long term dynamics of the system, the first perturbation labelled A causes the retiming of a pulse train and the second perturbation, labelled B fails to regenerate in the system. Whereas in panel (b), when the system is perturbed (labelled C) during its transient phase, it is possible to access the five pulse solution.

Figure 2.12 presents experimental results that highlight the importance of perturbation timing on the long term dynamics of the system. Subplot (a) shows how a perturbation with an appropriate timing and amplitude can trigger a second pulse train thus switching from 1-pulse to 2-pulse regime. This is in excellent qualitative agreement with fig. 2.10 (b) in the basins of attraction. A more intricate case is shown in figure 2.12 (b). As mentioned before, we time our perturbations using an external trigger. In this case, the external trigger sends two trigger signals, each approximately 40 ns wide and 3 μs apart. The first trigger signal generates two perturbations (due to the relative timing between the trigger and the pulse train from the Ti:Sa mode-locked laser). These two perturbations are 12.28 ns apart which is approximately $1.5 \times \tau$. These perturbations then generate

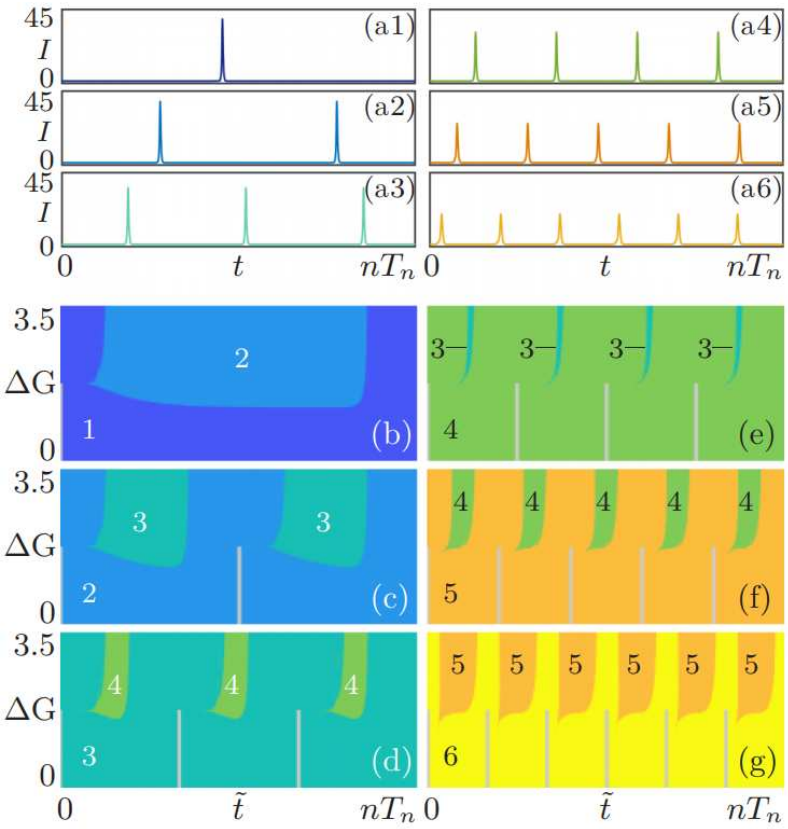


Figure 2.10: Multistability and basins of attraction
(a1-a6) Represent the intensity time series of the stable periodic pulsing regimes of Eq. 2.1 represented over n periods T , with n the number of pulses in the span of the delay τ. (b–g) Depict the basins of attraction of 2.1 in the plane of timing $\tilde{t}$ and amplitude ΔG of a gain perturbation, when one (b) to six (g) equidistant pulses initially exist in the external cavity. The color represents the number of pulses observed in the long-term in the external cavity (panels (a1 - a6) indicate the color code), and the vertical gray lines indicate the timing of the pre-existing pulses.

two pulse trains which by the virtue of periodicity in τ are almost equally spaced (τ/2 apart in pseudo-space). This corresponds to region 2 in the figure. The second trigger signal, which is approximately 360 round trips later, generates two more pulse trains by the same mechanism. The time between trigger pulses and τ are not commensurate which manifests as the relative timing difference between the pulse trains starting from B and C and the two preexisting pulse trains. Please note that the pulse trains starting from B and C are also approximately $\tau/2$ apart. In the short evolution after the second perturbation sequence, there are four pulse trains in the cavity (denoted by the region 4). The perturbation labelled as B is timed appropriately and triggers a pulse train whereas the perturbation labelled as C is sent very close to an existing pulse train and it only

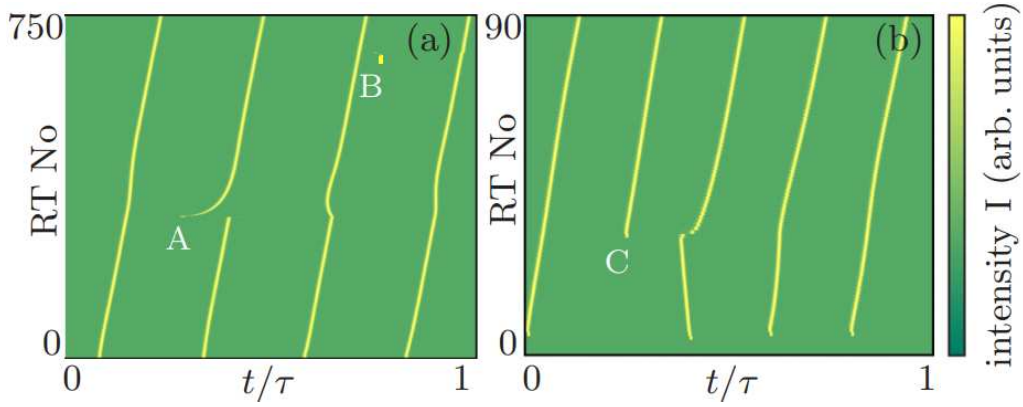


Figure 2.11: Simulated time traces showing transients and multistability
(a) Shows the effect of two perturbations labeled as A and B starting from a four-pulse solution.
(b) Shows a five-pulse solution reached by perturbing the system during a transient.

triggers a transient pulse train that is extinguished in a few hundred roundtrips. Thus the system settles to a three pulse regime denoted by region 3. As predicted by the basins of attraction analysis, these results indicate that accessing $n+1$ pulse solutions from n pulse solution becomes harder with increasing n.

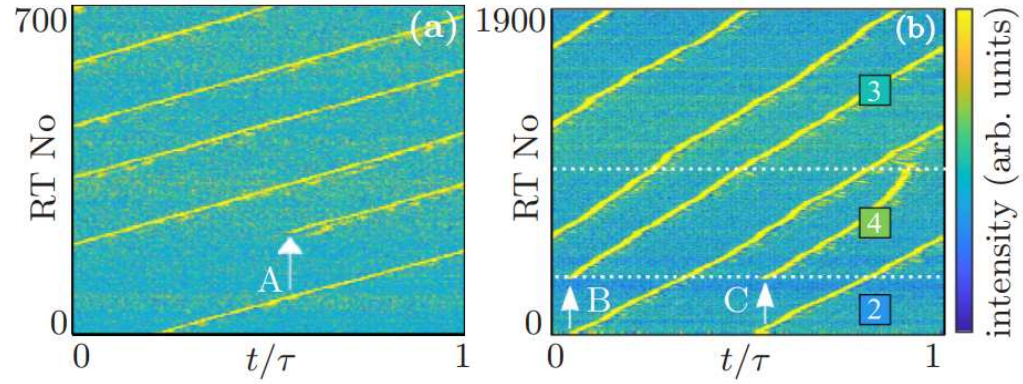


Figure 2.12: Experimental time traces showing transients and multistability
The figure shows pseudospace representation of time-traces showing response to external perturbations for one (panel (a)) and two pulse (panel (b)) regime. The feedback delay was 8.2 ns.

The same experiments can be carried out with coherent perturbations (at the cavity resonance wavelength) which corresponds to the perturbation on the intensity variable I instead of on the gain variable G in model 2.1. The basins were also mapped numerically with coherent perturbations ΔI. Apart from differences observed mainly in the finer details of the basins boundaries, the structure of the basins of attraction is qualitatively as

those shown in Figures 2.10 (b-g). This strongly indicates that the strength and timing of the perturbation is more important as compared to the perturbation variable.

2.4.2 Symmetry-broken pulsing patterns - Faster gain

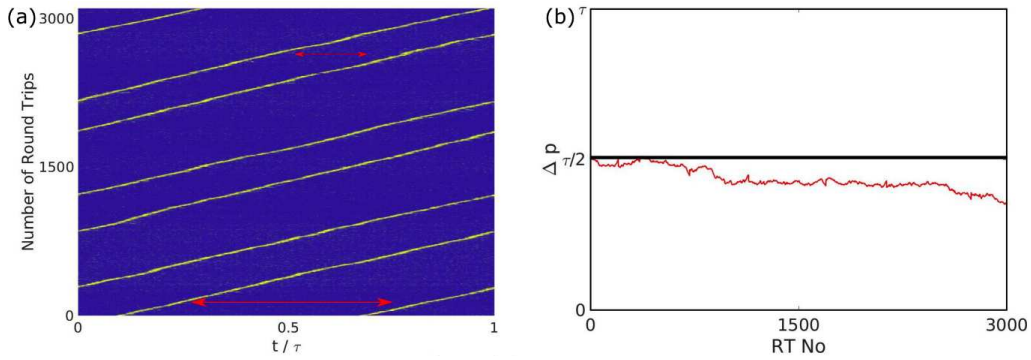


Figure 2.13: Stable non-equidistant pulse trains

Subplots (a) shows two coexisting pulse trains in the cavity. These two pulse trains appear to be equally spaced in the beginning but with the increasing number of round trips they seem to experience an attractive interaction and stabilize at non-equidistant values. Subplot (b) Plot the pulse-to-pulse timing Δp versus the roundtrip number (denoted by the red arrow in subplot (a)). We observe that while the two pulses initially have a distance of $\Delta \mathrm{p} = \tau/2$ in the beginning, the pulses undergo a gradual attractive interaction decreasing the value of Δp with successive round trips while experiencing some fluctuations that can be attributed to pump noise. The feedback time $\tau = 8.2$ ns in the experiment.

In the previous section, we illustrated how the stable pulsing patterns are equidistant in the case of faster saturable absorber. However, experimentally we also observe the contrary (shown in figure 2.13). In this figure, we see how a non-equidistant pulse train is stable over several thousand round trips [79]. Such results can be explained via the inversion of time scales between the two quantum wells. To further our analysis, we consider the following parameters otherwise stated: $A = 2$, $B = 2$, $\gamma_G = 0.01$, $\gamma_Q = 0.055$, $s = 10$ and $\kappa = 0.2$. The feedback time τ is treated as a bifurcation parameter. In this parameter region we report pulse-timing symmetry-breaking phenomenon, where some of the n-pulse solutions described in the previous section destabilize resulting in a coexistence of equidistant and non-equidistant pulse trains in the feedback loop. Experimentally, the parameter regime of faster gain is selected by choosing a suitable micropillar laser on the chip. As described in the fabrication of the micropillars, the etch quality determines

the non-radiative recombination rate; higher surface roughness leads to an increase in the number of defects which in turn increases the carrier recombination rate. The etch quality slightly differs throughout the sample resulting in different combinations of recombination rates and gives us access to a wide range of parameters on a single chip. The pump value can also change the effective recombination rate mainly in the gain section due to spontaneous emission. The effect is not accounted for in our model but can be easily introduced by an effective recombination rate combining all the effects. This way, either by tuning the pump and/or choosing a different micropillar we have have access to different parameters. A priori, both the sections have similar recombination rates thus these external parameters strongly influence the ratio of the recombination rates which is the important quantity for the symmetry breaking mechanism.

Figure 2.14(a) shows the one-parameter bifurcation diagram of the Yamada model with delay (Eq. 2.1) in the delay time τ, where the solutions are represented by the maximum intensity value attained (denoted by I_p). When τ is increased from 0, successive Hopf bifurcations (H, marked by blue dots at the bottom of subplot (a)) are encountered which leads to the coexistence of several periodic solutions. Each of the Hopf bifurcations correspond to a different frequency. The curve labeled as 1, corresponds to the fundamental solution with one pulse per feedback loop appears at $\tau = 51.7$ and is stable for all values of feedback delay. On the other hand, all the n-pulses solutions with n $\geq$ 2 emerge unstably from a Hopf bifurcation, subsequently stabilize in a torus bifurcation when τ increases, these solutions correspond to the periodic emissions of short pulses of light with periods close to sub-multiples of τ. On increasing τ further, these solutions destabilize through a second bifurcation. All these solutions coexist with the zero-intensity equilibrium solution (not shown in the bifurcation diagram) *id est* the laser off solution, which is stable over the entire range of τ.

Figure 2.14 subplot (b) presents the enlargement of the previous figure near the destabilizing bifurcations of the equidistant pulsing regimes with two to five pulses (points P, T_3, T_4, and T_5 respectively). The two pulse solution destabilizes at point P via a period doubling bifurcation. Prior to this bifurcation, the two equidistant pulses can be seen as a periodic solution with period $\tau/2$, following this bifurcation, the two pulses become

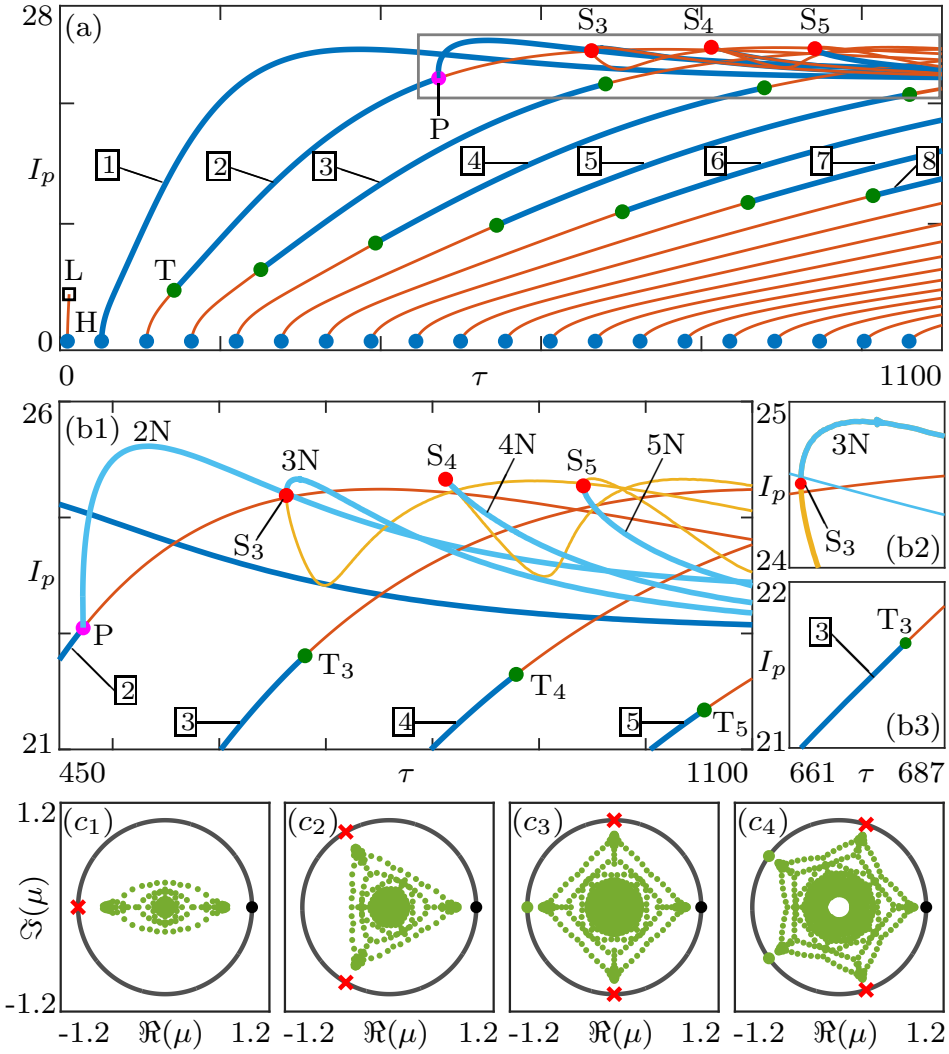


Figure 2.14: Bifurcation analysis for the Yamada model with faster gain

(a) Bifurcation diagram of 2.1, showing the pulse intensity I_p with respect to τ , with the number of pulses per feedback loop along each stable periodic solution branch. (b1) Enlargement of the framed area in (a), with further enlargments around point (b2) T_3 and (b3) S_3. Stable equidistant (E) and non-equidistant (N) pulse solutions are represented in dark and light blue, respectively, and unstable E and N solutions in dark and light orange, respectively. The dots indicate Hopf (H), torus (T), period doubling (P), saddle-node (S), and homoclinic (L) bifurcations.

non-equidistant in the feedback look and thus appear as a pulse-timing symmetry broken state with a period of τ. This non-equidistant state has two pulses with different amplitudes which is not represented in the Figure 2.14 (a) as it only plots the maximum value of I for every solution. After the period-doubling bifurcation at $\tau = 472$, one observes both a splitting of pulse amplitudes and timing. This this represented in the Figure 2.15. Panel (a1) shows the bifurcation point P and the subsequent peak values of both the pulses

in the symmetry broken state with increasing τ. Panel (a2) depicts the change in inter-pulse timing t_p with increasing τ. This amplitude and symmetry breaking of the relative pulse timings is due to the strong amplitude-time coupling of the system [80, 81]. The numerical integration of the system gives additional insight into the convergence of the symmetry broken state. Figure 2.15(b) shows the long term dynamics of the system with $\tau = 1000$ starting from an unstable equidistant two-pulse solution. A small perturbation is applied on the gain variable G which causes the system to begin converging into one of the two possible non-equidistant stable pulsing patterns: the first pulse timing interval decreases (2.15 b2) and the second pulse (highlighted in gray) converges towards a low amplitude state (Figure 2.15(b1)). When a different initial perturbation is applied by depleting G slightly (not shown here), the phase-shifted, symmetric version of this solution is obtained, with the first (green) and second (gray) pulses converging to the low-amplitude and high-amplitude states, respectively. Although this leads seemingly to the same long term dynamics, both of these different states occur, one being a phase-shifted version of the other. We also point out that the convergence is very slow and occurs over several thousand of delay times, showing that the stable non-equidistant solutions are only weakly attracting.

The three to five pulse solution undergo a different mechanism of destabilization, these points are marked as points T_3, T_4, and T_5 respectively. Following these points, a pair of stable and unstable solutions emerge from a saddle node bifurcation (these points are labelled as S_n for n=3, 4, and 5). For example, S_3 emerges at $\tau = 663$ for $n = 3$. The emerging periodic solutions following the saddle node bifurcation have a period close to τ as opposed to τ/n of the n-pulse solution before undergoing the torus bifurcation. Here, the solution following the saddle node bifurcation corresponds to the pulsing regime with n non-equidistant pulses of different amplitude in the feedback loop. At any time, τ, a vertical cut can give the number of coexisting pulsing solutions in the external cavity. Based on the subplots (b2) and (b3) we see that there is a small region of overlap in the value of τ that can support both equidistant and non-equidistant pulsing solution for $n = 2$. This can also be seem more evidently for $n \geq 3$ in subplot (b1). This sort of coexistence gives rise to an increased level of multistability in the system.

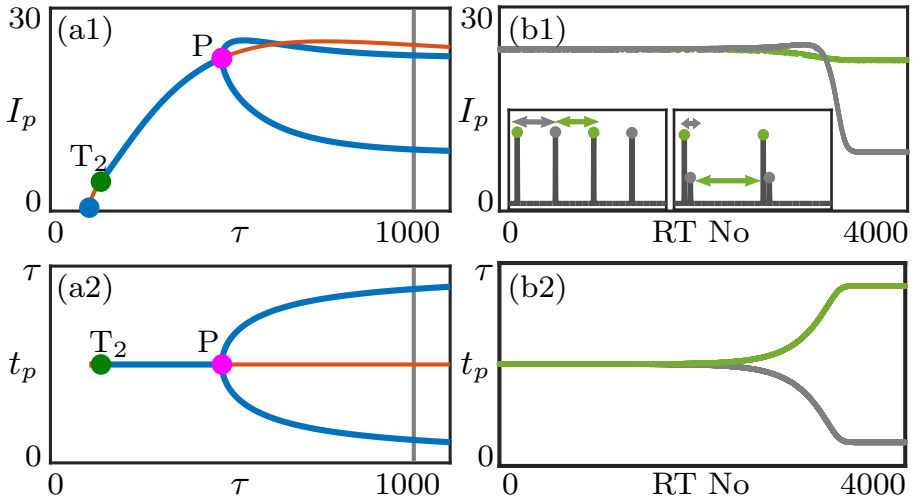


Figure 2.15: Pulse-timing symmetry breaking with two pulses
(a1) Maximum I_p of pulse intensity and (a2) relative inter pulse timings t_p along the branches of two equidistant and two non-equidistant pulses, with respect to τ . Stable and unstable solutions are represented in blue and red, respectively. (b) Simulation of 2.1 for $\tau = 1000$ [gray lines in panel (a)] with initial condition very near the (unstable) two-pulse solution, showing the long-term evolution of (b1) I_p and of (b2) t_p. The insets in (b1) show the intensity time series during the two first and two last roundtrips through the feedback loop; the dots and arrows indicate the amplitudes and relative timings as represented in (b1) and (b2), respectively

A different way of presenting this information is by expressing it in terms of the temporal pulsing patterns, this is shown in figure 2.16. Subplot (a) shows the representative time traces of one feedback delay τ for all the coexisting solutions for $\tau = 1000$. We observe the coexistence of 1-pulse solution, 2,3,4,5-pulse non-equidistant solution (panel a2-a5) and 5,6,7-pulse equidistant solution (panel a5-a7). Increasing the feedback delay progressively destabilizes the equidistant solutions and stabilizes non-equidistant solutions. Qualitatively, for a large value of τ, there is a coexistence of n-pulse solutions among which smaller values of n correspond to non-equidistant solutions and the larger values of n correspond to equidistant solutions. In Figure 2.16, all the periodic solutions with 1-7 pulses per feedback loop coexist but the ones with 2-5 pulses underwent the resonance tongues transition and thus are non-equidistant solutions. The stability regions in the (τ,κ) plane of 1-8 pulse equidistant and non-equidistant solutions are presented in Figure 2.16(b). Here the regions of stability of the non-equidistant pulsing solutions are resonance tongues bounded by saddle-node bifurcations. The stability regions of both types of solutions extend over large areas of the (τ,κ) plane and show a high degree of

multi-stability (panels b1-b8 for 1-8 pulse solutions). The long-term convergence to one or other pulsing solution depends critically on the chosen initial conditions. For $n \geq 2$ we observe a finite overlap between the two categories of solutions signifying that both these solutions are stable for the same parameters over some area on the parameter plane. As shown in Figure 2.14(b1), this results from the fact that the saddle node bifurcation occurs (at points S_n) slightly before the n-pulse solutions destabilize at the torus bifurcation points T_n. Hence, in these parameter regions of τ, one observes the n-pulse equidistant or non-equidistant solution depending on the initial conditions. For the chosen parameters in Figure 2.16, the 5-pulse solution displays this phenomenon (panels (a5) and (b5)).

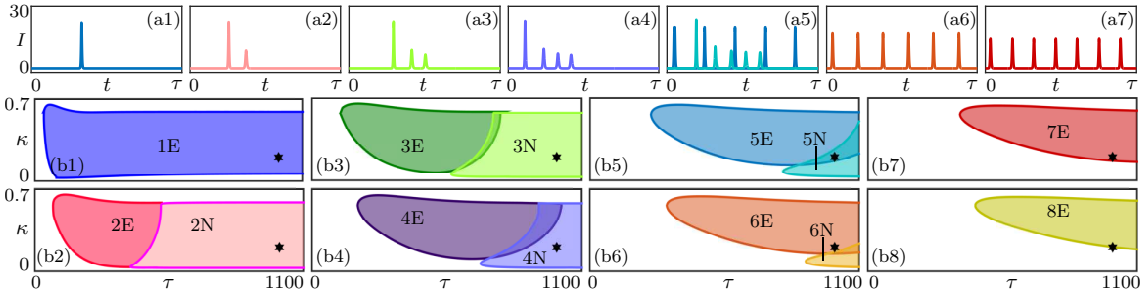


Figure 2.16: Coexistence of solutions in Yamada model with faster gain
(a) Intensity profiles of coexisting periodic solutions of (1), for $\tau = 1000$ and $\kappa = 0.2$. (b) Regions of stability, in the (τ,κ)-plane of feedback parameters, of the families of equidistant (E) and non-equidistant (N) periodic solutions with one to eight pulses per feedback loop. The number of pulses is indicated in the colored regions, and the star indicates the parameter point $(\tau, \kappa) = (1000, 0.2)$ of the time series in panels (a).

Experimental results shown in figure 2.17 (a and b) show the time traces recorded after two and three external perturbations respectively. In both the panels, the external perturbation are timed to generate responses close to the asymptotic 2 and 3 pulse equidistant solution. The panels (a1) and (b1) show the time trace recorded for two roundtrips soon after the initial perturbation, however in the long term, the timings between the consecutive pulses converge to unequal values (panels (a2) and (b2)) showing clearly that the equidistant solution is not stable.

Figure 2.18 shows the same type of convergence over several thousands of roundtrips but for different initial conditions, the colour coding of the inter pulse distance is the same in Figure 2.17. In figure 2.18 (a) and (b), the system is initialized with two and three

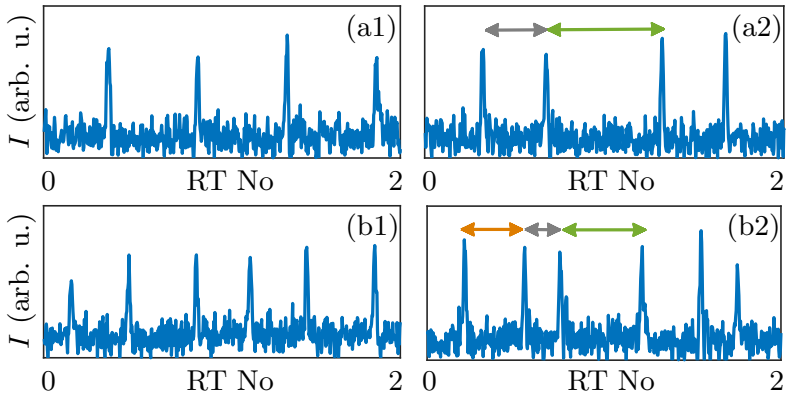


Figure 2.17: Time trace showing convergence from a symmetrical state to a symmetry broken state

(a1) and (b1) Show a representative time traces shortly two and three perturbations respectively. (a2) and (b2) Show a representative time trace after a large number of roundtrips in the cavity. The colour of the arrows corresponds to the inter-pulse interval shown in 2.18. The feedback time in the experiment was 8.2 ns.

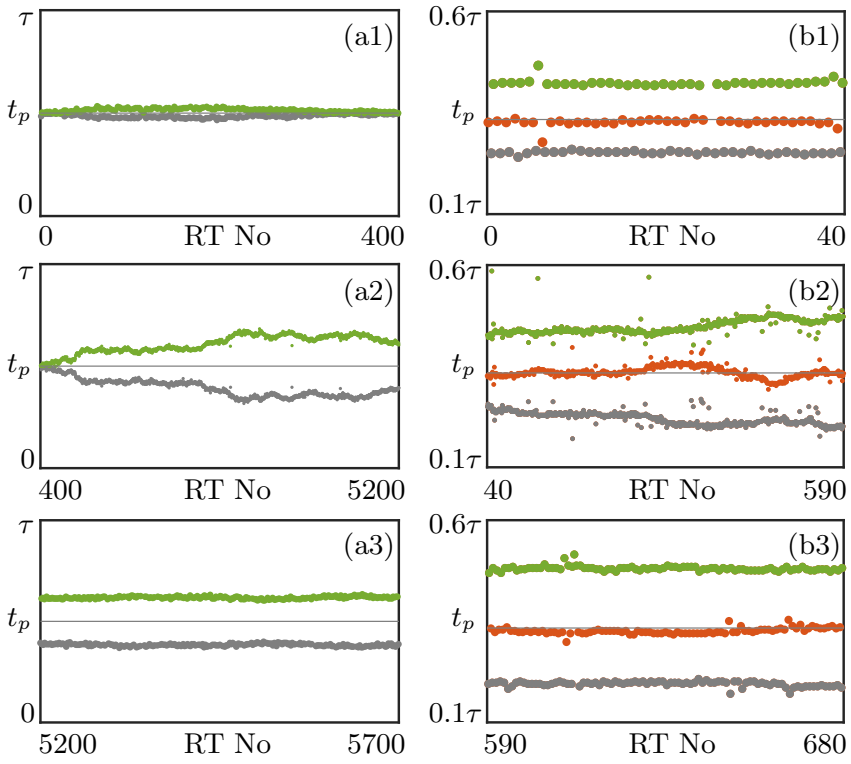


Figure 2.18: Evolution of inter-spike interval during the convergence process from a symmetrical state to a symmetry broken state

Subplots (a) and (b) Show the convergence of a two-pulse and a three-pulse solution from a symmetrical state to a symmetry broken state. The roundtrip number 0 corresponds to the time trace just after the perturbation.

equidistant pulses per roundtrip respectively. This is done using suitable external perturbations. We observe that the pulse-timing information is preserved in the short term [64, 80]. On the other hand, in the long term, the system slowly converges towards non-equidistant pulsing patterns with well-defined and different inter-pulse relative timings. These inter-pulse timings then stay very stable over a large number of roundtrips. It was not possible to monitor the amplitude difference in the final state due to the limited signal to noise ratio –the emitted pulse energy is on the order of $\approx$ 100 fJ. Since it is observed in Figure 2.15, that a small difference in the amplitude is associated with a large inter-pulse interval, we conclude that the amplitude measurements are not relevant in the experimental recordings. Overall, the experimental observations show excellent agreement with the dynamics predicted by the bifurcation analysis of the model. They demonstrate multistability between the experimental regimes with two and three non-equidistant pulses. Moreover, the quasiperodic regime corresponding to unlocked dynamics on an invariant torus are not observed, in good agreement with the theoretical predictions of very large locking regions in the parameter space as seen in Figure 2.16.

2.5 Polarization Dynamics

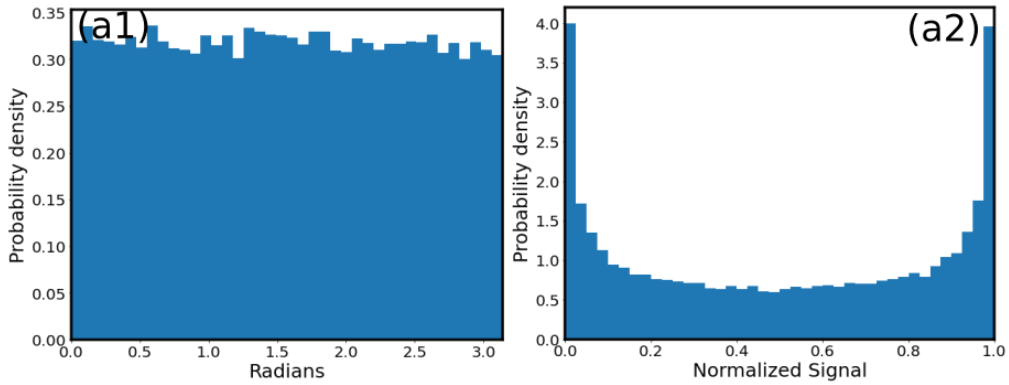


Figure 2.19: Probability density function of choice of random angle and the resulting intensity measurement

Subplot (a1) shows the realization of choosing 50000 random angles from a uniform distribution between 0 and π and subplot (a2) shows the resulting intensity measurement

In the previous sections we illustrated the intensity and temporal dynamics of the micropillar laser with feedback. Polarization dynamics in such a system can demonstrate a

variety of effects. In ref. [82], it was shown that a VCSEL with co- and cross-polarized feedback while emitting a constant output intensity, displays vectorial dissipative solitons following a periodic rotations in the linear polarization state. These rotations manifest as short pulses when detected through a polarizer. Other works on commercial VCSEL systems were used to realize neuron like functionality based on polarization switching [83, 38]. For reservoir computing applications, polarization dynamics in VCSELs has been shown to improve the memory and computing capacity of reservoir computing applications [84, 29, 85]. Polarization dynamics in VCSEL with SA has been theoretically investigated in ref. [86] and inhibition dynamics have been demonstrated. In this section, we present some experimental results on the polarization dynamics of micropillar lasers with and without feedback.

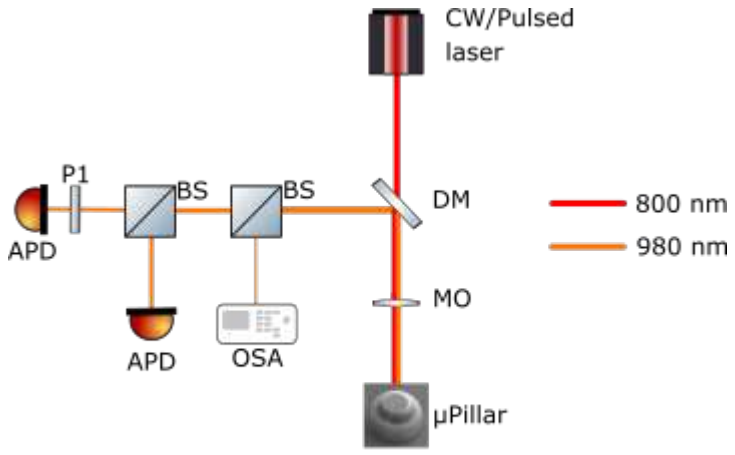


Figure 2.20: Experimental setup for measuring linearly polarized component of the excitable response oriented along a particular axis
DM: Dichroic mirror, BS: Beam splitter, MO: Microscope objective, APD: Avalanche photodiode, μPillar: Micropillar laser, P1: Linear polarizer oriented along the laboratory reference frame, OSA: Optical spectrum analyzer

We first present the polarization dynamics of a single micropillar laser without feedback to characterize its response and then demonstrate dynamics of such a system with feedback. As a first order approximation, neglecting internal dichroism and birefrengence, one can assume that in a circular cavity every polarization mode in π radians is degenerate. Thus if such a cavity is in the excitable regime and if an external incoherent perturbation produces a response, its polarization would be randomly drawn from an uniform proba-

bility distribution. This random choice of electric field oriented at θ radians when filtered through a linear polarizer and measured with a detector yields in the probability density function shown in figure 2.19. Subplot (a1) shows the probability density function of choosing a random angle from a uniform distribution and subplot (a2) shows the probability density function of detecting a particular intensity value. The intensity PDF is due to the $\cos^2(\theta)$ response, θ being the angle between the polarizer orientation and the electric field orientation. When the cavity under consideration is a semiconductor laser cavity the dichroism and birefringence play an important role and the degeneracy in modes is removed: certain orientations are preferred over the others. To characterize the emission properties further we use the experimental setup shown in figure 2.20. The CW laser is used to maintain the micropillar laser in the excitable regime and the mode locked laser is used to perturb the micropillar over the excitable threshold. Each perturbation from the Ti:Sa laser can be seen as independent excitation as the repetition rate of the mode locked laser (12.28 ns) is much greater than the internal time scale of the micropillar laser. Statistics from the measured intensity can then be used to infer information about the probability of emission along different orientations.

The results for such a measurement are shown in figures 2.21 and 2.22 for a circular micropillar and an elliptical[7] micropillar respectively. A sequence of above threshold perturbations result in a sequence of excitable responses, these responses are split using a 50:50 (R:T) beam splitter and sent to two detectors. On one of the detector arms, we insert a linear polarizer oriented along the laboratory frame of reference. In the figures, the central plot shows the joint probability distribution of the polarization sensitive ($p(I_p)$) and polarization insensitive ($p(I)$) intensity measurement. The individual probability densities of the polarization sensitive and insensitive intensity measurements are plotted on the sides. It is evident from the figures that the polarization of the electric field is not drawn from a uniform distribution. This effect is even more pronounced in the case of elliptical micropillars which show that the polarization can be pinned to a particular orientation. A complete characterization of the emission of the micropillar lasers would require a simultaneous measurement using four detectors (as shown in figure 2.23) to obtain a

[7]Elliptical micropillars are fabricated by altering the shape of the mask to obtain elliptically shaped micropillars. The dimensions of the large and small axis are 5 μm and 4 μm respectively.

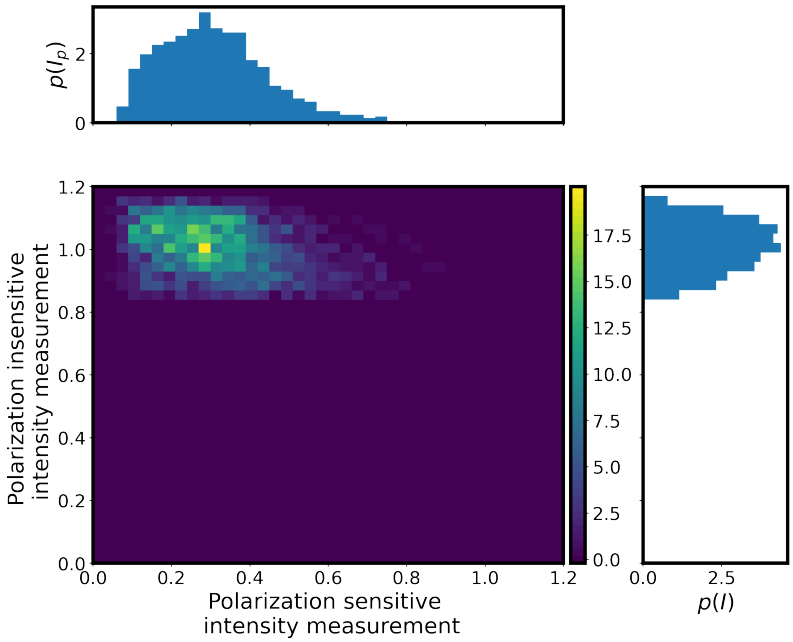


Figure 2.21: Probability density function of polarization insensitive intensity measurement and polarization sensitive intensity measurement in a circular micropillar

The central plot displays the joint distribution of the values of the polarization sensitive and insensitive intensity measurement. The individual probability densities are plotted on the sides. The polarizer was oriented along the laboratory reference frame.

probabilistic description of the Stokes parameters of the laser. We are currently limited by the signal to noise ratio in the experiment to carry out this full characterization. The preliminary results from the elliptical micropillars seem to be encouraging to demonstrate inhibition dynamics which could be useful for certain computational schemes.

Having demonstrated that an incoherent perturbation leads to a probabilistic choice in the output polarization state, we proceeded to study the effect of coherent perturbation on such micropillars. Coherent perturbation scheme can be realized in two ways:

- Tuning the central wavelength of the mode locked laser to be in the cavity resonance of the micropillar. This can be very difficult to achieve since the cavity resonance is smaller than the free spectral range of the Ti:Sa laser (0.25 nm) which makes it compulsory to use thermal tuning of the laser cavity.

- Exciting the micropillar laser using an incoherent perturbation and then using the excitable response as a coherent perturbation.

The second method can be realized using an external cavity as demonstrated in this

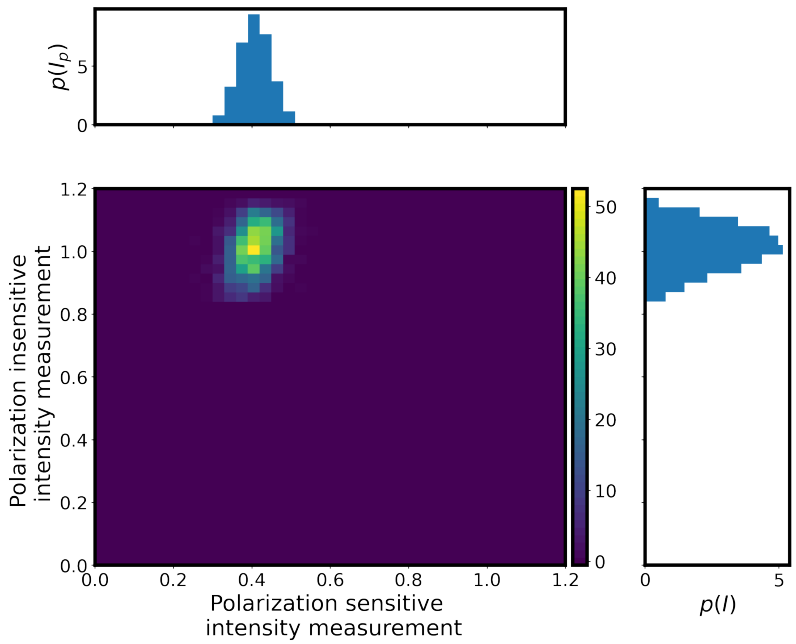


Figure 2.22: Probability density function of polarization insensitive intensity measurement and polarization sensitive intensity measurement in a elliptical micropillar

The central plot displays the joint distribution of the values of the polarization sensitive and insensitive intensity measurement. The individual probability densities are plotted on the sides. The polarizer was oriented along the laboratory reference frame.

chapter. The single incoherent perturbation now triggers a train of responses, each of which can be viewed as a coherent perturbation for the next response. We then perform a polarization sensitive and insensitive measurement to resolve the dynamics of such a pulse train. The experimental setup used is shown in figure 2.24. This experiment was only performed with circular micropillars due to experimental constraints. The time traces obtained from the two detectors reveal interesting polarization dynamics. The Fourier spectrum of the polarization sensitive and insensitive intensity measurements are shown in figure 2.25. The Fourier transform of the polarization insensitive measurement reveals only a fundamental peak at $f = 1/\tau$ where τ is the feedback time which is approximately 8.8 ns. The polarization sensitive measurement reveals sidebands around the fundamental peak. The inset shows how the modulation sidebands change with respect to the the bias pump. The modulation time is approximately 39.8 ns, 40.32 ns, and 41.15 ns for pump values of 590 mV, 597 mV, and 608 mV. The pump span was chosen such that the micropillar was still in the excitable regime. The modulation time undergoes a small change of approximately 1.4 ns for the large range of pump values. The pres-

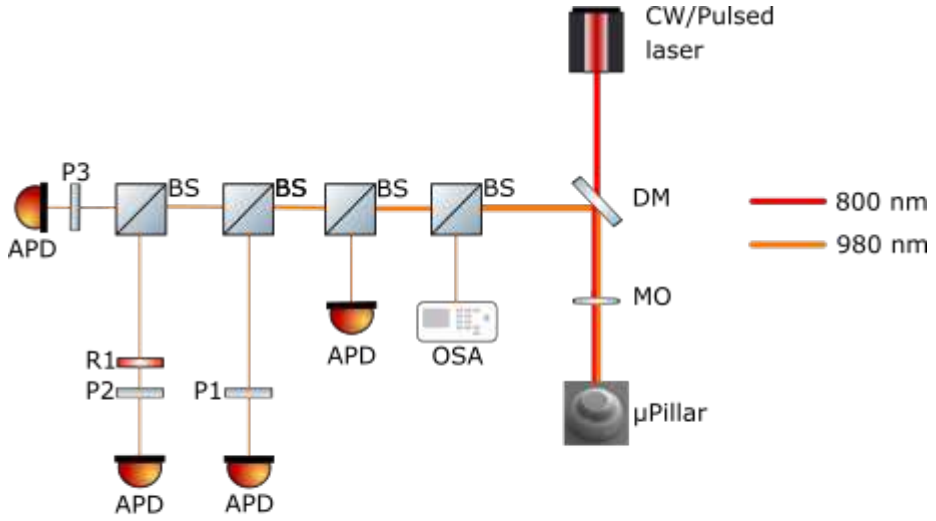


Figure 2.23: Experimental setup for complete characterization of polarization state of the excitable response

DM: Dichroic mirror, BS: Beam splitter, MO: Microscope objective, APD: Avalanche photodiode, μPillar: Micropillar laser, P1: Linear polarizer oriented along 0 radians, P2: Linear polarizer oriented along 0 radians, P3: Linear polarizer oriented along $\pi/4$ radians, R1: Quarter waveplate oriented along $\pi/4$ radians, OSA: Optical spectrum analyzer

ence of modulation sidebands in the polarization sensitive intensity measurement and their absence in the total intensity measurement is a clear signature of modulation of the orientation of the electric field of the excitable response rather than a modulation of the intensity. We are currently collaborating with the group of Bernd Krauskopf from the Dodd-Walls Center for Photonic and Quantum Technologies, The University of Auckland, New Zealand on developing a model to explain the measurements. The time scale of the modulation in polarization could stem from a temperature dependent phenomenon. However, we have preliminary results showing that such dynamics could be explained without the temperature being a required parameter. Additional studies would be needed to explain the physical origin of this phenomenon. Another physical reason for ruling out the temperature dependence would be due to the very small change in the modulation time observed over a large range of pump values. Since the pump has a direct impact on the temperature and there is no appreciable change in the modulation, we could indeed assume that the modulation stems from a more fundamental reason.

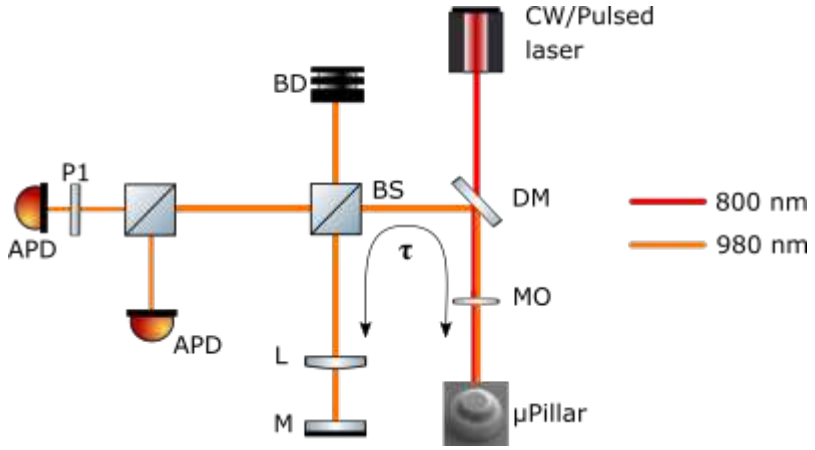


Figure 2.24: Experimental setup for measuring the linearly polarized component of the excitable response oriented along a particular axis for a micropillar laser with delayed optical feedback

DM: Dichroic mirror, BS: Beam splitter with 70/30 power split between reflected and transmitted path, MO: Microscope objective, APD: Avalanche photodiode, L: Lens with f = 5 cm, M: High reflectivity feedback mirror, BD: Beam dump, μPillar: Micropillar laser, τ : External cavity round trip time.

2.6 Discussion and Conclusions

The results presented in this chapter demonstrate the multistability of a micropillar laser with a time delayed feedback. The system, in the short term, demonstrates functionality as an optical buffer which can be addressed using single optical perturbations or by altering the bias pump. In the long-term the behaviour can be distinctly differentiated into two regimes. In the case of a faster gain region, it is shown that any initial pulsing pattern will converge to an equidistant pulsing pattern based on the initial conditions. It is also possible to perturb a system in its stable state and the long-term dynamics are then predicted by the basins of attraction. In the case of the a faster saturable absorber region, we observe a similar kind of convergence but to symmetry-broken pulse-timing states. Theoretically, our collaborators have shown that the stable n-pulse equidistant solutions destabilize via torus bifurcations. In the vicinity of the torus bifurcation, a saddle node bifurcation appears which stabilizes the symmetry broken states that are shown in the bifurcation diagram (fig. 2.14). Experimentally, these states have been observed in the case of 2 and 3 pulses in the external feedback loop. Physically, the strong amplitude-time coupling is responsible for the strong amplitude time coupling is responsible for the

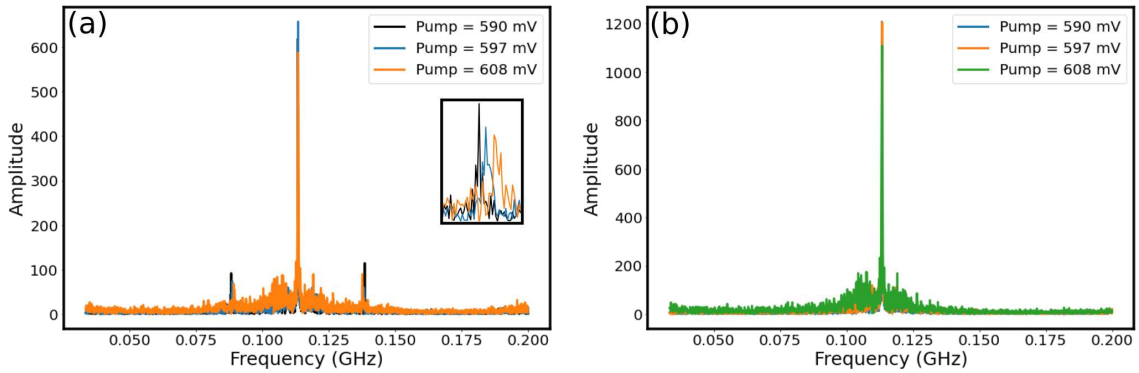


Figure 2.25: Fourier spectrum of polarization sensitive and insensitive intensity measurement in micropillar laser with feedback

Subplot (a) shows the fundamental peak corresponding to the feedback time and the side bands signifying a modulation in the electric field orientation. Further, the inset shows a slight change in the modulation frequency as a function of the bias pump. Subplot (b) shows that in a polarization insensitive intensity measurement the only observed peak in the Fourier spectrum corresponds to the feedback time.

occurrence of symmetry broken states. As the feedback time is increased, there is an increasing degree of multistability among symmetric and symmetry broken states.

The results presented here are quite generic in essence that they only require excitability and a time delayed feedback, thus their implications can exist beyond optics. Our results contribute towards the understanding different pulsing dynamics observed in other fields. Some notable references on related subjects are [87, 88] which demonstrated the effect of time delayed feedback on single limit cycle oscillators and [89, 90] which discusses effects of two limit cycle oscillators coupled through time delayed feedback. Since the system presented in this chapter is not a limit cycle oscillator, the oscillations exist due to the feedback and the time period is intimately linked to it, thus making the results novel. Based on recent results demonstrating a connection between temporal dissipative solitons in spatially extended systems and pulsing regimes in delay systems [81], possible connections might be made between non-equidistant pulsing regimes and soliton molecules as the former are bound states of pulses [91, 92].

Beyond their fundamental interest for study of nonlinear dynamics of delay systems, the results presented can contribute to the realization of optical computing schemes such as reservoir computing relying on the large phase space of delay systems [93, 94, 95, 30, 26], optical buffers [37, 59, 96] and content addressable memories [97]. The physical

meaning of a content addressable memory as described by Hopfield is an appropriate phase space flow of the state of a system. The interest of such systems is that it can retrieve a stored input based on a partial and/or error ridden input. In the context of the results presented in this chapter, the stored patterns correspond to the stored memories in the system and the input states correspond to partial matches to the stored memories. Based on the initial excitation pattern, the system converges to one of the asymptotic pulsing patterns which can be also viewed as the minimization of a certain energy defined for the system. The energy landscape ideally contains a local minima for every stored pattern and based on the initial conditions, convergence to the nearest local minima. Another direction of research would be to compute using attractors created by several coupled micropillars with delayed optical feedback. The hardware setup would be similar to the proposed arrays of coupled micropillars via a diffractive optical element as in [30] but would have a fundamental difference in the sense, the information would be written into the system at time $t = 0$ and allowed to evolve within the system and might converge to a possible attractor. This would then be an interesting demonstration of having all optical computing without any optical to electronic conversion layers.

Finally, in the last section of the chapter, we demonstrate the polarization dynamics of a single micropillar (circular and elliptical) lasers with incoherent perturbation and single circular micropillar with coherent perturbations which are realized with the help of delayed optical feedback. We observe that the in the case of the incoherent perturbations, the circular micropillar produces a response with a varied orientation of the electric field. On the other hand, the geometrical modification in the elliptical micropillars enable the pinning of the polarization with in a small range. The results from the coherent excitation (via delayed optical feedback) reveal a modulation in the polarization which is absent in the intensity measurement. The reason behind this is currently not completely understood.

Chapter 3

Computing with integrated micropillar lasers

In this chapter we present mainly numerical results on computing using either spatially coupled micropillar lasers or using an ensemble of uncoupled micropillars. We first describe the method of spatial coupling and the model. We show the propagation of excitation and how they can be used to build information processing circuits: OR, AND gates and a temporal pattern recognition circuit. We discuss the implementation of such circuits experimentally. We also show a design for realizing an on-chip excitable oscillator. Finally, we present numerical results on computing using an ensemble of uncoupled micropillars using a particular example of temporal code called rank order coding

3.1 Spatially coupled micropillars

2-D lattices of nodes can be built and coupled via out-of-plane elements [30, 31, 98]. The resulting setups can be large and complex to build while giving impressive computational ability. An alternative way to scale up the number of coupled nodes is to design integrated circuits and introduce on chip coupling. In the following sections we present the various functioning blocks such as tunable coupling, delay lines, and threshold detectors essential for building circuits using micropillar lasers. Coupled excitable nodes have been theoretically and experimentally studied in the past. Experimental studies include

semiconductor quantum-dot lasers using free-space coupling [58] and excitable microring lasers coupled through waveguides [36]. Theoretical studies have been carried out in coupled waveguides and optically injected microdisk lasers [99]. We propose in the following a different approach and study spatially coupled micropillars in order to analyze some computing properties of such networks. First we introduce the notion of coupling and denote the diffractive coupling constant with the term κ. The system of ordinary differential equations then used to model such evanescently coupled micropillars will consist of $3 \times n$ equations where n is the number of micropillars (Eqs. 3.1) [47, 100, 101, 102].

$$\begin{aligned} \dot{E}_j &= ((1 - i\alpha)G_j - (1 - i\beta)Q_j - 1)E_j + i\kappa_{jm}\, E^m \\ \dot{G}_j &= \gamma_G(A - G_j(1 + |E_j|^2)) \\ \dot{Q}_j &= \gamma_Q(B - Q_j(1 + s|E_j|^2)) \end{aligned} \tag{3.1}$$

Where the dynamical variables E_j, G_j and Q_j are the electric field, rescaled gain and absorption in the j^{th} micropillar. The time in these equations is rescaled to the cavity photon lifetime τ_p which is taken to be on the order of several picoseconds ($\approx$ 1-2 ps). The coupling term $\kappa_{jm}\, E^m$ follows the Einstein summation notation where the contributions from all the nodes is summed up. The coupling constant κ is real to have a purely diffractive coupling. When the system under consideration is a chain of micropillars with uniform diffractive coupling, κ can be reduced to a real number and the diffractive coupling for the j^{th} node can be written as $i\kappa(E_{j-1} + E_{j+1})$. Non-radiative carrier recombination rates and linewidth enhancement factors for the gain and absorber are γ_G, γ_Q, α and β respectively. A denotes the pump and B is the non-saturable losses. The saturation parameter is $s = \frac{a_Q \gamma_G}{a_G \gamma_Q}$ where $a_{G,Q}$ is the differential gain and absorption respectively; it controls the characteristic response time of the system. The laser threshold for a single micropillar as introduced in the first chapter is $A_{th} = 1 + B$. For a coupled system in the excitable regime, for $A \lesssim A_{th}$, an initial perturbation above the excitable threshold can propagate to the neighbouring micropillars in the saltatory propagation regime provided $\kappa \ll 1$. In our analysis κ is normalized using the cavity photon lifetime and $\kappa \ll 1$ condition ensures the formation of the excitable response before coupling to the neighbouring micropillars.

This model is also only valid for small coupling strengths when coupled mode theory is valid.

To quantify the coupling coefficient, we simulate two coupled micropillars. When two micropillars are placed in the vicinity of each other, the eigenmodes undergo mode splitting. The fundamental mode split can be expressed as a symmetric and anti-symmetric mode. Based on this splitting, one can estimate the coupling of energy from one micropillar to the other and assign a characteristic coupling time. This time when normalized with respect to the cavity photon life time is the value of κ considered here. The detailed calculations behind this are illustrated in the chapter 2 of ref. [44]. Figure 3.1 shows how κ is dependent on the distance between the center of two micropillars of radius $r = 2.5\mu m$.

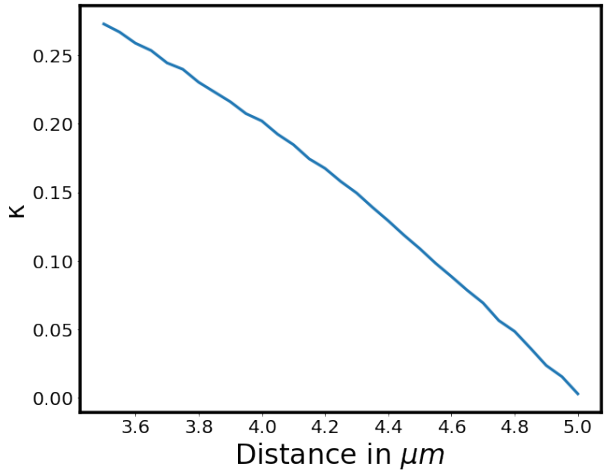


Figure 3.1: Coupling constant κ as a function of distance between two micropillars of radius $r = 2.5\mu m$

3.1.1 Numerical Simulations

Having outlined the prerequisites, we consider the following parameters for simulating a chain of micropillars (each micropillar is coupled to its neighbours): $\gamma_G = \gamma_Q = 0.001$, $B = 2$, $s = 10$, $\alpha = 2$, $\beta = 0$. The parameters A and κ are varied in a region to illustrate different behaviour. These parameters are compatible with the semiconductor parameters [103] and correspond to a non-radiative recombination time of 1-2 ns. The initial conditions are taken as $G_j(0) = A$, $Q_j(0) = B$, and $E_j(0) = \delta_{i0}F_0$ for all micropillars where δ_{i0} is the Kronecker delta symbol. While the simulations are done for the chosen

parameters, the phenomenon is robust and is observable over a large range of parameters. We choose to number the micropillars as $0 \cdots$ N from left to right, then the leftmost micropillar is perturbed at time $t = 0$ using the initial condition $E_0 = 5$ which is enough to overcome the excitable threshold (for the given value of A) and elicit a response from the micropillar. As a result of the coupling, the excitable response pulse is coupled to the neighboring micropillar and since the coupling strength is sufficient, the excitable threshold is crossed again leading to another pulse. The excitation transfer continues giving rise to a solitonic and ballistic response. Importantly, the response is unidirectional: the excitation can only transfer to the nearest neighbour on the right, or more precisely to the unperturbed micropillar. This happens due to the refractory period of the micropillars; the refractory period is longer than the coupling time between micropillars which results in symmetry breaking of the system and unidirectional propagation of the signal.

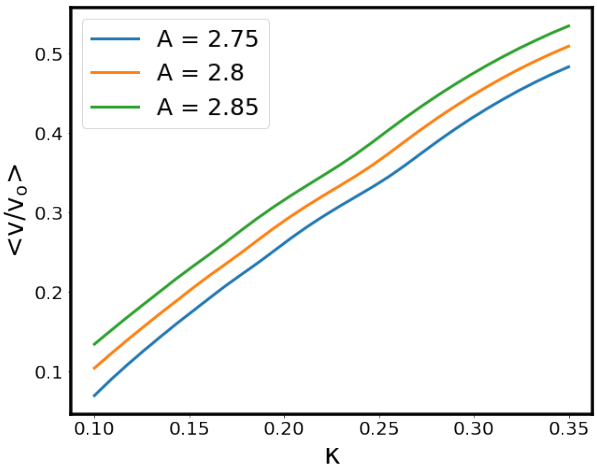


Figure 3.2: Average adimensional propagation speed of a signal $< v/v_o >$ (units of speed described in text) in a coupled micropillar chain for different values of pump A

The calculated mean speed of the pulse for A = 2.75, 2.8, and 2.85 are plotted in figure 3.2. The average propagation speed $< v/v_o >$ here is related to the time (scaled to the cavity photon lifetime) taken for the excitable pulse to propagate one adimensional unit of length with $v_o = 1/\tau_p$. It does not account for the physical distance between two micropillars which varies with the coupling strength κ (see fig. 3.1). The speed varies almost linearly with κ in a large window of laser parameters (only a small range of pump is shown here). The increase in κ directly translates to lower photon tunneling time be-

tween the cavities which increases the propagation speed. There is a slight deviation from linearity at higher values of coupling (> 0.3). In this region of high coupling, we observe an apparent saturation in the propagation speed (not shown in the plot). This has been identified in [101] as a bifurcation in the propagation mode of the excitation. The saturation of propagation speed can however be intuitively understood as follows: the coupling strength controls the time it takes for the energy to be coupled to the adjacent micropillar, when the excitable pulse amplitude is large and if the coupling strength is sufficient, the excitable threshold is easily reached. Hence the latency time does not play a large role and we expect a saturation in the propagation speed. For a low value of κ (typically < 0.1), no pulse can propagate. The numerically computed speed is typically on the order of a few hundreds of picoseconds for traversing 20 micropillars. This indeed represents a challenge experimentally because one needs a high time resolution to resolve the propagation from individual cavities; additional limitations are caused by the low emission intensity.

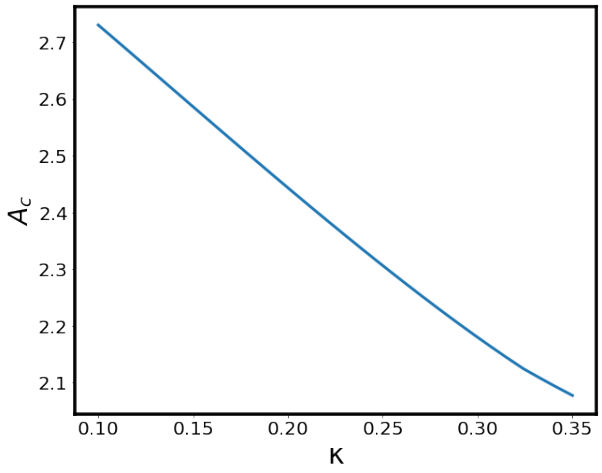


Figure 3.3: Critical pump A_c required for signal propagation for different coupling strengths κ

For a given coupling strength κ, there is a minimum value of the pump A below which the excitable propagation is inhibited. This is illustrated in figure 3.3. This critical value of pump decreases with increasing the coupling strength. The intuitive explanation of this is that small coupling strengths must be compensated by higher excitable pulse amplitude and this is possible by increasing the pump. This is consistent with the numerical and

analytic findings of refs. [49, 50], where it is shown that the maximum excitable pulse intensity scales as $\gamma_G^{-1} \times (A - 1 - ln(A))$. Having introduced the parameters required for signal propagation, figure 3.4 shows the signal propagation or lack thereof in two chains of twenty micropillar lasers with different coupling strengths. Here the pump is $A = 2.8$ and the other parameters are as mentioned before. A perturbation above excitable threshold is introduced in the leftmost cavity (0) in a chain of 20 coupled cavities. In subplot (a), the coupling constant $\kappa = 0.01$ is too small to induce propagation: only the first cavity fires an excitable spike. In subplot (b), the coupling strength $\kappa = 0.1$ is sufficient to induce stable saltatory propagation in the chain.

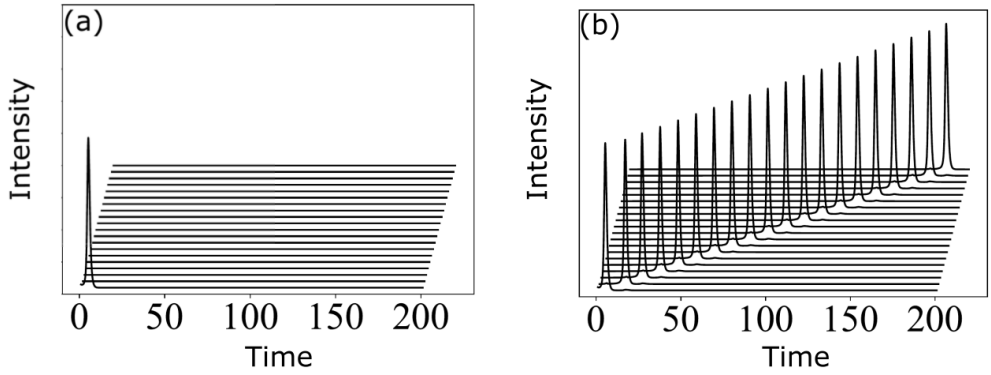


Figure 3.4: Saltatory propagation for different coupling strengths
In (a) and (b), the temporal response (normalized to cavity photon lifetime) of each cavity is plotted with an offset for clarity. The coupling constant in (a) is $\kappa = 0.01$ and in (b) is $\kappa = 0.1$. Adapted from [102]

3.1.2 Information processing

In an effort to build information processing circuits based on the understanding established in the previous section, we first illustrate the implementation of temporal AND and OR logic circuits. These logic gates are analogous to their static counter parts in terms of their truth table but depend on the relative timing of the excitable response from the micropillars. Out of deference to spiking neural networks, these excitable responses can also be called spikes[1]. Another way to view logic circuits can be as a classification problem. A visual schematic of the truth table of the OR, AND, and XOR logic circuits is shown

[1]The words spikes and excitable response are used interchangeably in the thesis, I have tried to add footnotes in a few places in order to avoid any confusion to the reader

in figure 3.5. Evidently, there exist several linear decision boundaries in the case of the OR and AND logic gate but a nonlinear decision boundary is required for the XOR operation. Like the journey from single layer Perceptrons [1] to modern neural networks is one of linear to nonlinear decision boundaries, we show how coupled micropillars can *learn*[2] linear decision boundaries and in principle be extended to nonlinear decision boundaries such as XOR logic [86, 104] and ring neural networks [105].

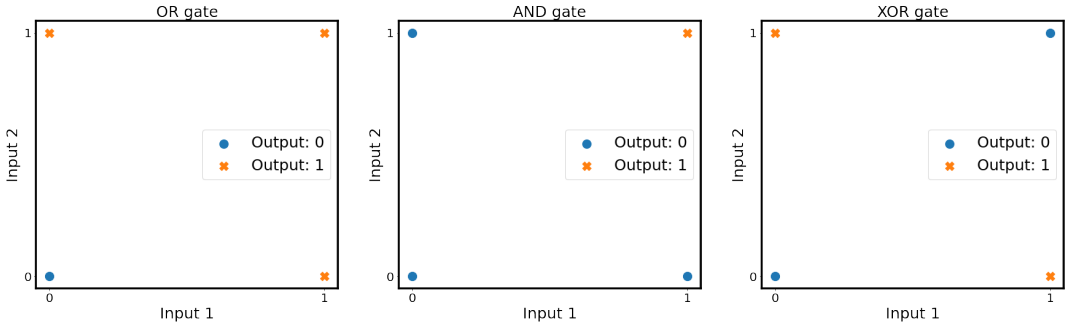


Figure 3.5: Truth table for OR, AND, and XOR logic

Linear decision boundary problems

Boolean logic operations with excitable systems have already been studied in models of dendritic spines [106], using chemical excitable waves and their collision properties in 2D media [107, 108] and in semiconductor media with excitable localized states [109, 110]. The first illustration is that of the OR logic operation. In these circuits, information is coded in the presence or absence of a spike and can be represented by two bits, 1 and 0. These circuits are also designed to be cascadable. We consider a chain with an odd number of micropillars ($n = 13$) for the main segment, to which is attached an additional segment in the center consisting of a single micropillar. The OR gate is illustrated in figure 3.6. The two input micropillars are A and B and the output micropillar is C. The above threshold perturbations are marked by arrows. Except in the case where no input is present, the output of the gate in C is always 1. With only one input, the response is trivially 1 as the single excitation (input at either at A or B) propagates throughout the structure like in the chain of micropillars. If both the inputs are 1, the pulses collide in the central micropillar

[2]The word 'learn' is used liberally and refers to the fact that the parameters are decided before hand and fixed for particular applications.

and excite the output micropillar C. Therefore, a spike emerges and can propagate to possibly another gate. This forms an OR gate. Note that where the two incoming pulses meet, the response latency time is smaller.

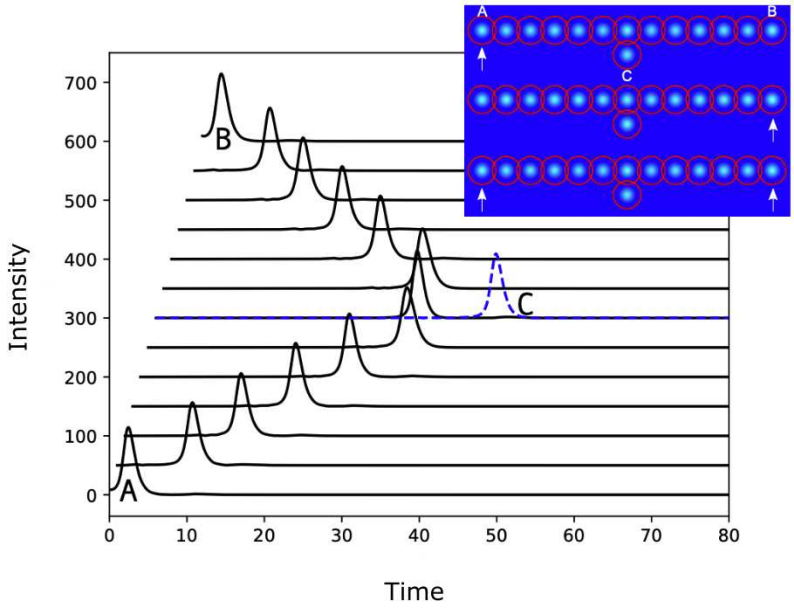


Figure 3.6: Numerical simulation of a spike-based OR gate circuit
Temporal responses for each cavity (offset for clarity) in the case of two perturbations in nodes A and B. The response in node C is plotted in dashed and blue. Inset: overlayed responses for inputs in nodes A, B, and both A and B. The cavities are pumped with $A = 2.74$. The coupling strength is $\kappa = 0.15$. Adapted from [102]

The AND gate is less trivial and requires to modify either the coupling or the pumping of the micropillars. Since the latter is easier to implement experimentally, we choose this option for the numerics. As was noticed earlier, there is a critical pump for the spike to propagate. We use this property to build the AND logic gate. A temporal AND gate can also be viewed as a coincidence detector gate. We consider the same chain as before, except the micropillars immediately next to the central cavity are now pumped with a lower value (see Fig. 3.7). The value is chosen to prevent the propagation of a single pulse. However, when the two inputs are present, the coupling of the two lower intensity pulses in the central micropillar is sufficient to cross the excitable threshold and a pulse is created in C. The lower pump is immediately translated into a larger spike latency time and a smaller response pulse. When these smaller pulses are coupled into the central micropillar, they can produce a response pulse in C which can be cascaded eventually to

other gates. The same phenomenon arises for a constant pump value of $A = 2.74$ and a change of the coupling constant of the nodes next to the vertex node, for values in the range $0.05 \lessapprox \kappa^{'} \lessapprox 0.1$. If $\kappa^{'} \lessapprox 0.05$, the pulse cannot propagate anymore and is stopped completely. If $\kappa^{'} \gtrapprox 0.1$, the gate transforms into an OR gate since a single pulse can propagate through the vertex.

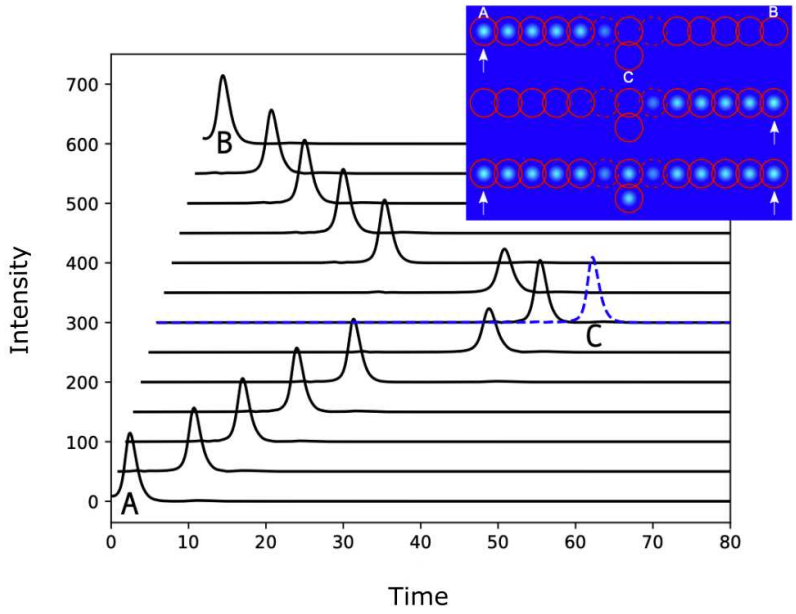


Figure 3.7: Numerical simulation of a spike-based AND gate circuit
Temporal responses for each cavity (offset for clarity) in the case of two perturbations in nodes A and B. The response in node C is plotted in dashed and blue. Inset: overlayed responses for inputs in nodes A, B, and both A and B. The cavities are pumped with $A = 2.74$ except the one marked by dashed circles with $A = 2.45$. The coupling strength is $\kappa = 0.15$. Adapted from [102]

The last example is that of a temporal pattern recognition circuit (a particular example using delay lines and three excitable nodes was shown in [46]). This circuit is capable of recognizing when two consecutive spikes in an input sequence are separated by Δt ($= 515.5$ shown in figure 3.8). The input signal is sent to micropillar A and the output is recorded from micropillar D, the response is recorded at several micropillars in the circuit to demonstrate the operation of the temporal recognition. The input spike propagates through the upper arm (see inset in Fig. 7). At the first crossing, the signal is split in two parts. The lower arm implements a delay line whose delay is equal to Δt. Both arms meet in a node structure similar to the previously shown AND gate. Every time the

temporal pattern in input is recognized, a spike is emitted at the output. To speed-up the computation time and reduce the size of the delay line we have accelerated the carrier dynamics in the simulation by taking $\gamma_G = \gamma_Q = 0.01$ in Eqs. 3.1. The input pattern recognized corresponds essentially to the propagation delay of the pulse in the lower arm. Generalization of the circuit to the recognition of more complex temporal patterns can be done via integrating several delay lines. However, the refractory time of the nodes sets the lower limit of the detectable separation between two input spikes.

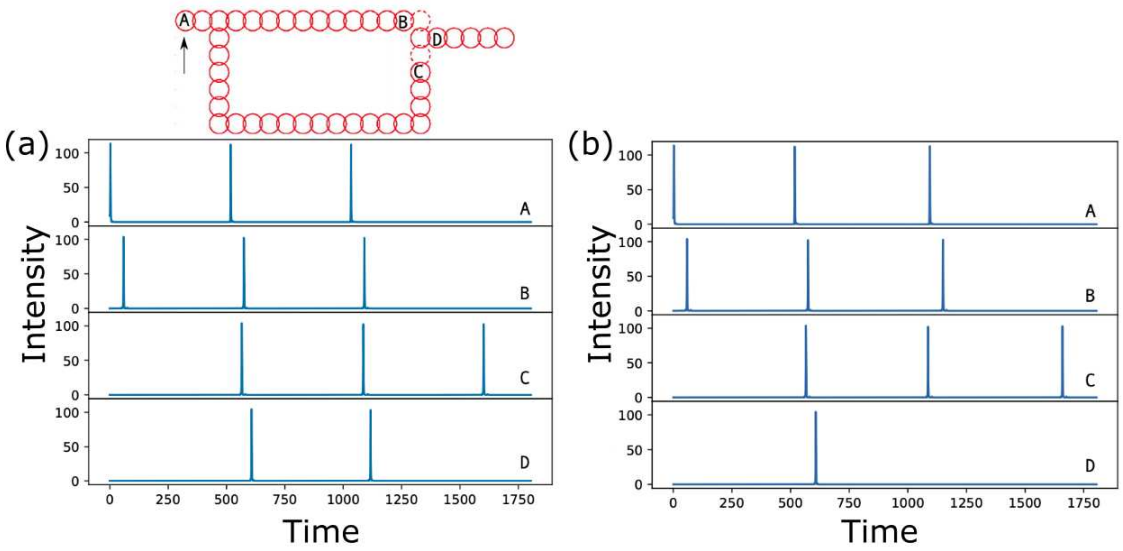


Figure 3.8: Spike based temporal pattern recognition
A signal is sent to node A and recorded in nodes A, B, C, and D. In node D, the output pulse signals the presence of specific temporal pattern: two spikes separated by $\Delta t = 515.5$. In (a) the input signal consists of three spikes separated by Δt. In (b) the input signal consists of two first spikes separated by Δt and a third spike separated by $\Delta t' = 575.5$. In inset is shown the micropillar circuit considered. The micropillar are uniformly pumped except the two pillars with dashed circles (same parameters as in Fig. 3.7 except $\gamma_G = \gamma_Q = 0.01$). Inset: schematic of the circuit used (the lower arm length has been reduced for clarity). The total number of nodes is 105. Adapted from [102]

While the topology of the spike pattern decoder is fixed by fabrication, the Δt value can be adjusted by modulating the pump intensity thus making it possible to recognize various temporal patterns using a single fabricated structure. The functionality illustrated here can also be used to build on chip oscillators which generate a spike train akin to the external cavity systems illustrated in chapter 2. Note that this functionality is not trivial, if one considers e.g. a ring topology for the coupled pillars, one cannot have self sustained oscillations: any perturbation introduced in the ring gives rise to two counter propagation

excitations which would eventually collide and suppress the propagation. A schematic structure is illustrated in figure 3.9. The circuit in terms of pumping or coupling is identical to the temporal AND logic. The micropillars marked in solid line are pumped higher (or have a higher coupling) as compared to the micropillars in dashed lines. An initial spike / perturbation is sent to node A, this spike is then split into three possible paths labelled as Paths 1, 2, 3, and 4. The three paths can be alternatively expressed in the order of the nodes traversed. The nodes traversed in paths 1 and 2 are A-C-B-A and A-D-B-A respectively. The nodes traversed in path 3 and 4 are A-B-C-A and A-B-D-A respectively. The time taken for traversing the paths A-B-C and A-B-D is shorter than time taken by paths A-C-B and A-D-B due to the design. As illustrated in the temporal AND logic, the signal cannot conduct via path 3 and 4 due to lower pumping (or coupling). The paths 1 and 2 are symmetric and thus due to this reason, the initial perturbation at node A results in a spike at nodes C and D at the same time. This makes the conduction in the full path possible. The same phenomenon reoccurs every time node A spikes resulting in a on chip oscillator with the period fixed by the length of paths 1 and 2, the global pump, and coupling constant. In the absence of the modified pump or coupling on paths A-B-C and A-B-D, the first perturbation at node A would split into all the four paths and conduct successfully along each of them. This would then lead to two collisions along paths 1,3 and paths 2,4 stopping the signal propagation. This circuit would thus no longer serve as an on-chip oscillator.

Non-linear decision boundary problems

Non-linear decision boundary problems can be approached in two ways, (a) Training a non-linear decision boundary using non-linear activation functions or (b) Adding features to increase the dimensionality such that the problem requires a linear decision boundary in higher dimensions. To illustrate the example of the latter, consider the XOR logic problem with the truth table as follows:

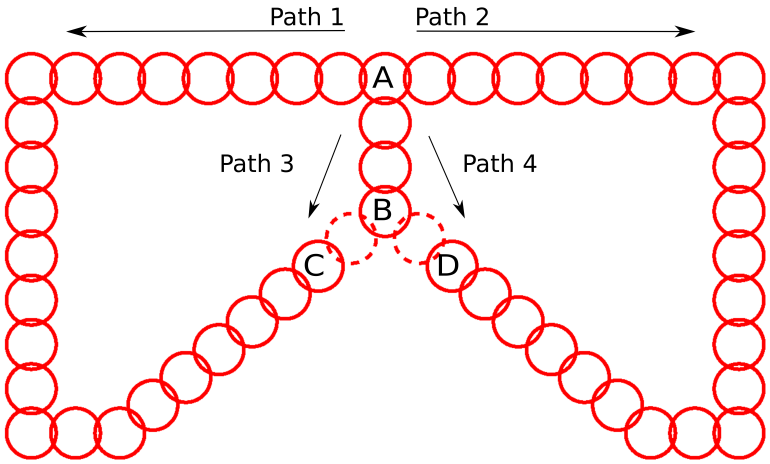


Figure 3.9: On chip oscillator
A signal is sent to node A which then splits into three paths, Path 1, 2 and 3 as marked in the figure. Paths 1 and 2 (of nodes A-C-B and A-D-B respectively) are viable paths for signal conduction whereas Path 3 (of nodes A-B-C or A-B-D) is not. This results in the performance if this as an on chip oscillator. See text for additional details.

Input 1	Input 2	Output
0	0	0
0	1	1
1	0	1
1	1	0

One can add an extra dimension using a logic operation such as $\mathrm{Input1} == \mathrm{Input2}$ and the problem with the following truth table them becomes linearly separable in three dimensions:

Input 1 == Input 2	Input 1	Input 2	Output
1	0	0	0
0	0	1	1
0	1	0	1
1	1	1	0

This approach is essentially one of the motivations behind reservoir computing, a

reservoir computer generates features and increases the dimensionality of the problem such that it can then become linearly separable in higher dimensions.

An optical XOR gate can be shown using the polarization degree of freedom [86] in a non integrated scheme. It has been shown recently in the group that it is also possible to obtain this logical operation in evanescently coupled micropillars.

Lastly, coupled micropillars can be used as generic nonlinear devices in schemes of reservoir computing by combining several on-chip oscillators and delay lines. Such circuits can then in principle be used to learn *arbitrary* decision boundaries.

3.1.3 Towards experimental realization

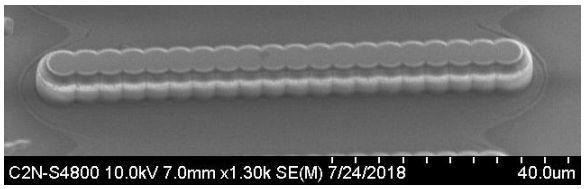


Figure 3.10: SEM image of 1-D chain of micropillar lasers

Having illustrated the results from numerics on coupled micropillar, it is essential to demonstrate that they can be implemented in actual devices. In figure 3.10, we see how we can fabricate coupled micropillars based on the same fabrication recipe for the micropillars. The distance between the center of micropillars is tuned to achieve a range of coupling constants.

In figure 3.11, we demonstrate the fabrication of a structure which can act as an AND logical gate. As indicated on the image, the input node is labeled with a blue circle and the output node is in black. To ensure that both arms are in phase, the path following the input node is split in two arms of equal lengths, realizing a type of interferometer. The circles labeled in red are the ones which are either less pumped (shown here) or less coupled to the output gate. The modulation in pump can be achieved by the means of a spatial light modulator (SLM) in the pump beam path. This modulation changes the arrival time of the spikes at the summation region. The output node only spikes if the change in time is within the duration supporting temporal summation ($\approx$ 600 ps) illustrated in the first chapter. The SLM can be used to imprint a phase profile on the input pump beam

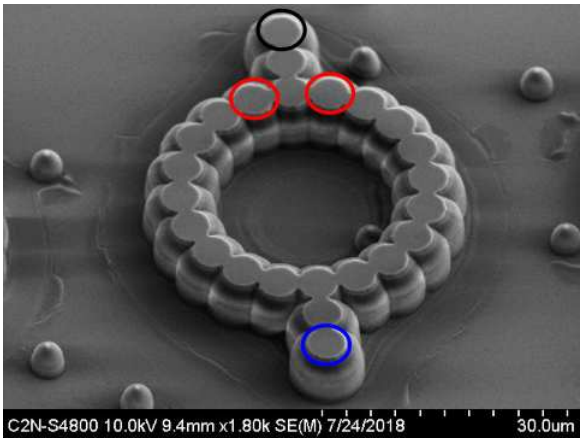


Figure 3.11: SEM image of coincidence detector / AND logic gate
In the figure here, the micropillar labeled with the blue circle is used as the input and the micropillar labeled with the black circle is used as the output. The two micropillars labeled with the red circles is where either the pump (shown here) or coupling can be altered to achieve the coincidence detector / AND logic gate.

which when focused generates a desired point spread function (PSF). Such problems to obtain a desired PSF have been extensively studied and one of the popular algorithms for phase retrieval is the Gerchberg-Saxton algorithm [111] which is an iterative procedure. This algorithm assumes a plane wave is incident on the SLM surface and only modifies the phase leaving the intensity constant. Several modification exist for tailoring to specific application of an experiment.

3.2 Brain inspired computing using an ensemble of micropillars

In this last section of the chapter, we will demonstrate a computing scheme based on rank-order coding implemented using micropillar lasers. While there are many paradigms of brain-inspired computing (some of which are introduced in the introduction), this section will focus on neural coding schemes. Neural coding schemes refer to plausible ways spikes or electrical impulses can contain information and perform computation. While this is still an active area of research and there is no definitive answer about the scheme used by biology, there are a few well known contenders. The first of which is rate coding, this is a relatively simple coding scheme and the information here is coded on the aver-

age number of spikes per unit time. One can easily see how such a scheme while easy to implement can be very inefficient. There have been experimental results suggesting some primitive computations occur at much faster timescales than enabled by rate order coding thus, there must be more efficient schemes [112]. Ref. [113] illustrates some elementary computations to compare three schemes *videlicet* rank-order coding, time coding, binary coding, and rate coding. The setting is as follows: Let there be N neurons, the temporal resolution of the measurement apparatus be t and the total acquisition time be T. The amount of maximum information (in bits) in binary coding (1 when a neuron fires and 0 otherwise) is $\log_2(2^{\mathrm{N}})$. If we were to discard the information identifying which neuron fired and only consider the number of neurons that fired, we obtain the maximum information to be $\log_2(\mathrm{N}+1)$ (the N is for the number of neurons and the additional 1 when no neuron fires) for rate coding. Time coding relies on precise timing of the spikes from each neuron, given the temporal resolution and the acquisition time, the spike from each neuron can be classified into T/t slots, generalizing this to N neurons results in the information capacity of $\log_2((\mathrm{T}/\mathrm{t})^{\mathrm{N}})$. The last case does not account for the time of each neuron to spike, just their relative order. N neurons can be ranked in $N!$ ways and the resulting in the information capacity of $\log_2(\mathrm{N!})$. Figure 3.12 shows how the maximum information scales in each of these cases. As we can see, time and rank order coding perform significantly better as the number of neurons increases. While these calculations describe the scaling of maximum possible information, specialized mechanisms (in biology) or hardware (in engineering) can be needed to extract pertinent information. A more practical calculation would be one that weights the maximum information against the degree of sophistication required to extract information. For several reasons, it turns out that rank-order coding is considerably simple to implement and thus we choose this scheme to implement computation using micropillars.

For demonstrating learning and computation, we choose the digits dataset from scikit-learn consisting of handwritten digits. Each digit is an 8×8 image with pixel intensities between 0 and 16. There are 1797 images with almost equal number of images in each class. The Yamada model with incoherent excitation is used in this illustration. We use the model equations introduced in chapter 1 (Eq. 1.1):

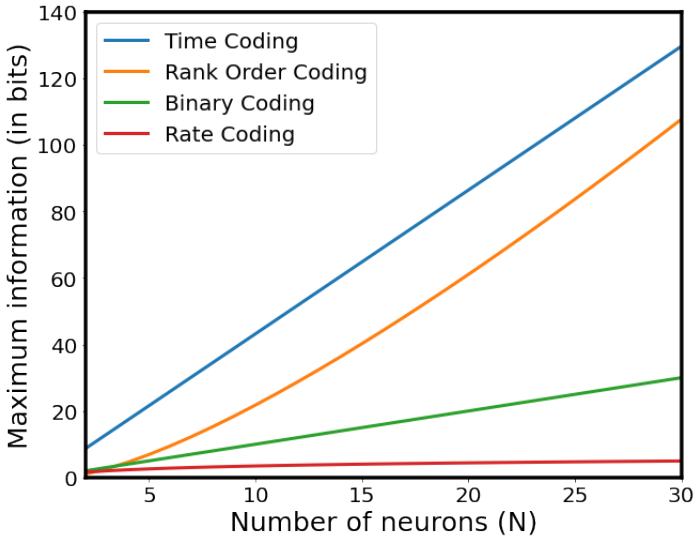


Figure 3.12: Maximum possible information in different coding schemes
The information in time-coding is calculated using $T/t = 20$

$$\begin{aligned} \dot{G} &= \gamma_G(A - G - GI) \\ \dot{Q} &= \gamma_Q(B - Q - sQI) \\ \dot{I} &= (G - Q - 1)I + \beta_{sp}(G + \eta_1)^2 \end{aligned} \tag{3.2}$$

The parameters used for the simulation are: $\gamma_G = \gamma_Q = 0.01$, $s = 10$, $B = 2.2$, $\beta_s p = 10^{-5}$ and $\eta = 1.6$. The choice of the pump A is explained in the next section.

The training procedure for rank-order coding is very minimal. We determine a weight vector for each digit by choosing a few digits (≈ 20) at random from each class and compute a pixel wise average. The result of this average is seen in figure 3.13.

The average values are then sorted in the descending order and each pixel is assigned a rank, $\mathrm{rank}_{\mathrm{px}}$, where px is the pixel number. The weight then associated to each pixel (or rank) is $W(px) = D^{rank_{px}}$, where D is a constant < 1. We choose $D = 0.95$. This introduces a decay in the weights. The resultant weights obtained are shown in figure 3.14.

For the testing or inference phase, each pixel is sent to an independent micropillar as a perturbation on the gain variable at time $t = 0$. As illustrated in the first chapter, the micropillar exhibits spike latency which is a function of the bias pump and the perturbation

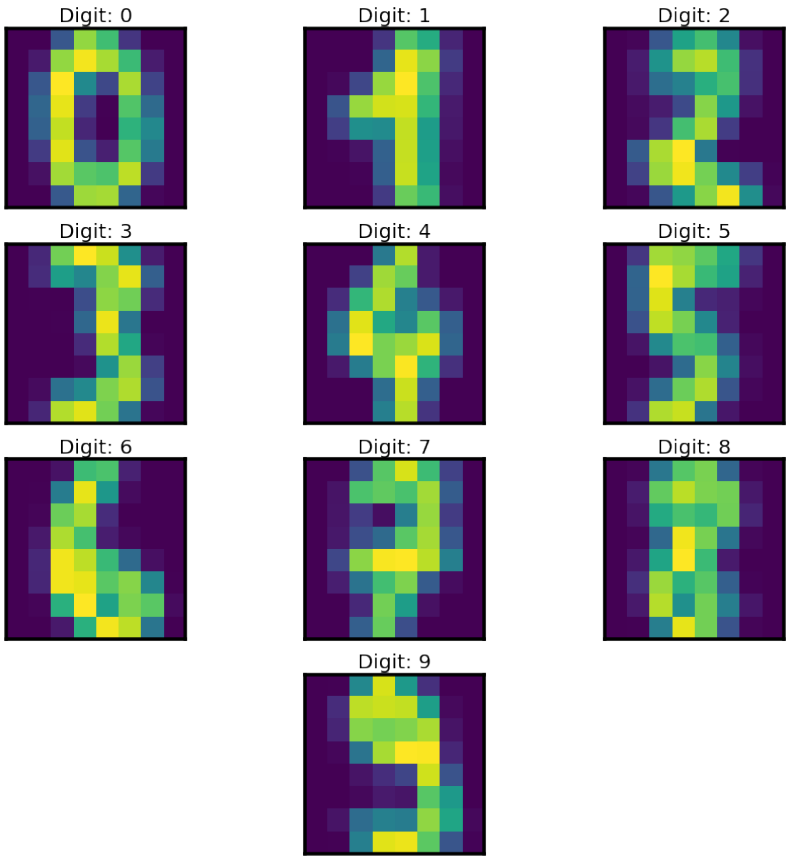


Figure 3.13: Pixel wise average computer for digits of each class (brighter is higher)

strength. This relationship for various values of pump is shown in figure 3.15. This plot helps us select a good operating point which demonstrates a large range of nonlinear spike latency for the micropillar. Once the bias pump is chosen, the handwritten digit data is rescaled to be between the excitable threshold and 1.05[3]. Thus every pixel is converted into a latency time t_{resp}. As a convention we take $t_{resp} \to \infty$ if a particular micropillar does not generate a response. Thus for each digit, we obtain 64 pulses at different times. We then rank the order of these pulses such that the shortest time t_{resp} is ranked 1 and $t_{resp} \to \infty$ is ranked ∞ in an order vector [4]. Once we have the weight vectors (from training) and an order vector for the test case, we compute the activations for every digit class as follows:

[3]This value chosen here is just to avoid being in the saturated region of the curve. It doesn't play a very critical role in the simulations as we use float64 numbers but this would be very critical in presence of noise or in experiments.

[4]Please note that if a micropillar does not spike, it is assigned $t_{resp} \to \infty$ and ranked ∞ and not 64, this is chosen as a convention

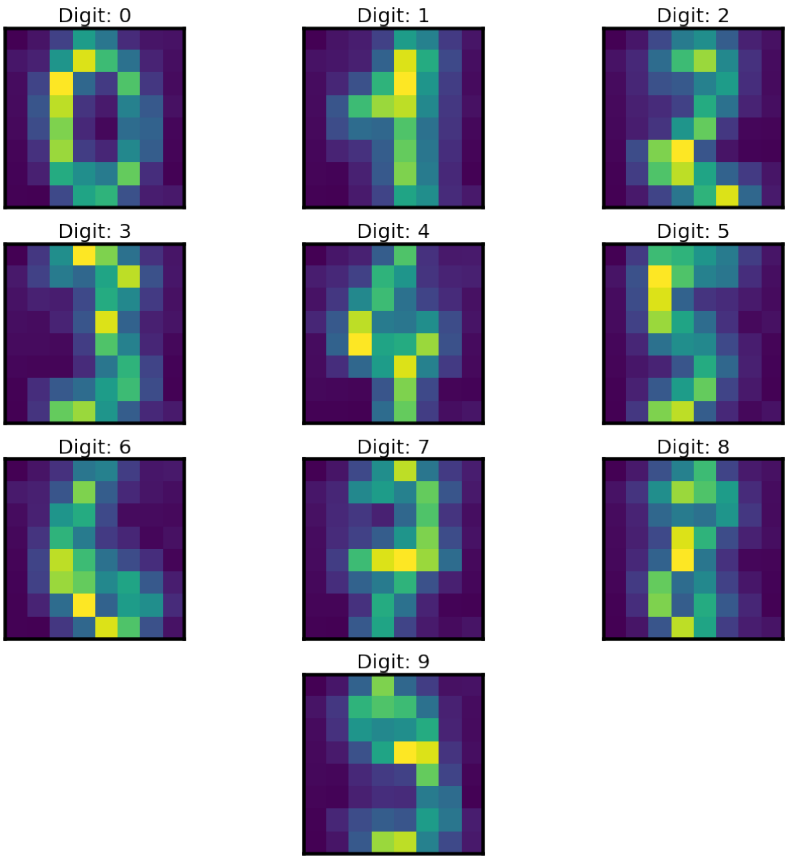


Figure 3.14: Weights for rank-order coding (brighter is higher)

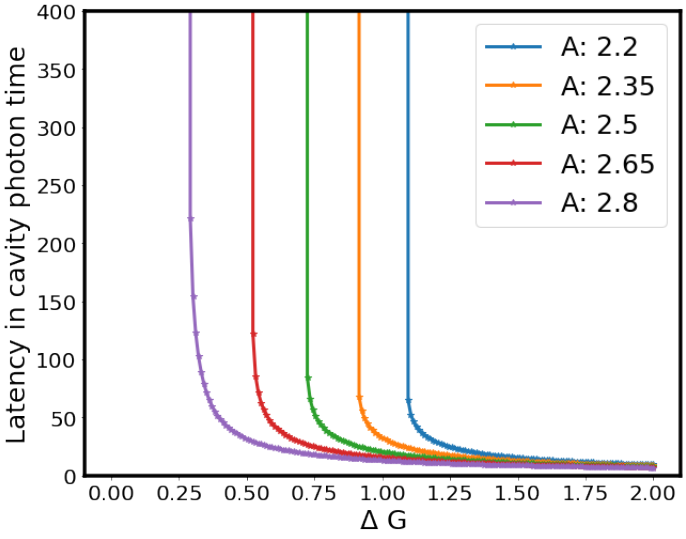


Figure 3.15: Spike time latency in micropillar laser

$$Activation(i) = \sum_{j=1}^{64} M^{order(a_j)} \times w(i,j) \tag{3.3}$$

Here, the index i between 0 and 9 corresponds to the number of test classes, the summation over j calculates the contribution from each pixel factoring in its rank and the corresponding weight. $M < 1$ is chosen by convention to introduce a decay in the contribution of micropillars with a high spike latency to the total activation. $M = 0.95$ is chosen in our case. For each test case we compute the 10 corresponding activations and the class with the maximum activation is taken as the predicted class. In doing so, we obtain a test accuracy of $\approx 74\%$. The breakdown with their respective classes are as follows:

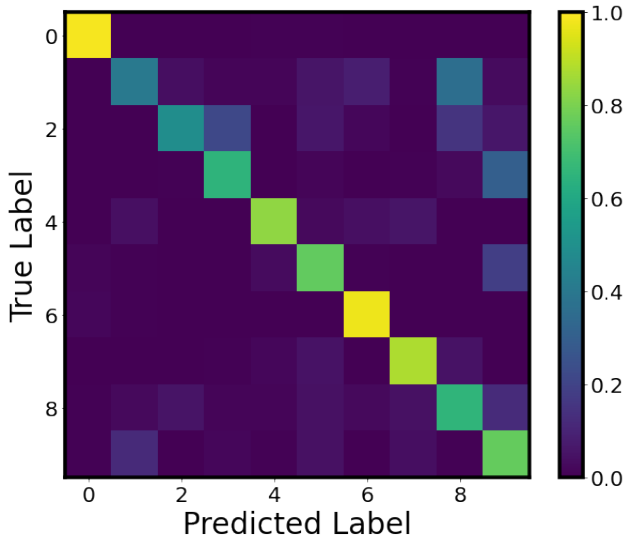


Figure 3.16: Confusion matrix for prediction using rank order coding

Class	Accuracy
0	0.987
1	0.402
2	0.490
3	0.650
4	0.835
5	0.759
6	0.975
7	0.876
8	0.653
9	0.764

The confusion matrix is shown in figure 3.16. The prediction results on some digits are better than others. One potential way to correct this is to identify a refined set of weights which better captures the digit. It is an open question and something the group will work on in the near future. The results presented do not compare to the near perfect accuracy obtained by the state of the art neural networks and machine learning techniques but present a novel form of computing using temporal information. We strongly believe that the accuracy obtained can be increased using optimal training routines. To the best of our knowledge, this is the first demonstration of computing using rank order coding on a photonic platform. Current research is underway in our group in the ANR Anaconda project to use more advanced neural coding schemes and build photonic hardware based on delay and spatially coupled nodes.

Chapter 4

Precursor aided prediction of extreme events in an extended microcavity laser

In this chapter we introduce extreme events formation in a quasi-1D semiconductor system and we demonstrate and discuss the prediction of such events. The prediction is carried out on experimental data displaying spatiotemporal chaos by using various machine learning techniques based on precursors identified using statistical analysis and information theory. Our analysis is based on a partial knowledge of the history of the spatiotemporal laser intensity. Thus, rather than attempting to predict the full spatiotemporal field, we attempt to predict the category of future events based on the knowledge of a finite history of the system, building on the identification of spatiotemporal regions carrying the most possible useful information. We show that we are able to obtain good results on the cross-prediction of extremes, based on the knowledge of the history at two spatially disconnected regions, on par with the best predictions using the most correlated regions of space.

4.1 Extreme Events: A historical perspective

Extreme events can be defined as follows [114]: a time recording of the dynamics of a macroscopic system maintained out of equilibrium may consist of well-defined periods where a relevant variable undergoes small variations around a well-defined level defined by its long-time average, with the occasional occurrences of abrupt excursions to values that differ significantly from the average level; such events are called extreme events. In recent years, significant efforts have been made to understand extreme events. These events were first observed at sea and for centuries were merely considered a myth despite anecdotal evidence as linear models used by researchers to predict wave amplitudes ruled out existence of such large waves. With the dawn of the modern industrial era and steel-hulled ships the probability of surviving encounters with rogue waves increased and with it the eye-witness records. It was a Scottish oceanographer named Laurence Draper who worked on analysis of waveform data and wrote the first scientific article on the subject in 1964 [115]. It was several decades before an experimental conformation of this phenomenon. On January, 1, 1995, an oil drilling platform located off the coast of Norway measured a wave measuring 85 feet which was appreciably outside the predicted models; at the time it was estimated that a 64 foot wave would occur approximately once every 10,000 years. The scientists from the oil rig later published [116] that such rogue waves were far from a rare phenomenon and occurred with a much higher probability. Rogue waves and their prediction is an active area of research and they can be observed in many systems with non-negligible probabilities [117, 118, 119, 120]. Scientists have also demonstrated how nonlinear interactions between waves can cause the existence of *super rogue waves* from calm and apparently safe sea states. These waves are caused by the nonlinear focusing of wave amplitudes of nearby waves into one large and short lived event [121]. Nonlinear systems are of particular interest as they can show the requisite extreme sensitivity to initial conditions.

4.2 Extreme events in optics

Extreme events in optics are characterized by a rare and intense pulse in the intensity domain. The study of these optical extreme events has been originally motivated by their analog in hydrodynamics because physically some conservative models of wave propagation in optics and deep water waves can share a common description by the the nonlinear Schrodinger equation [122]. The interest of optics stems from the fast timescales enabling large statistics to be recorded in a short amount of time and with tunable nonlinearities. Thus, many different physical situations can be tested and we can expect to draw general conclusions from extreme events found in optics when identified with well defined physical mechanisms. Most of the studies in this context have taken place in optical fibers where the interplay of nonlinearity, dispersion, and noise can generate extreme events [123, 124, 125, 126]. Extreme events in conservative systems are often associated with the merging dynamics of coherent structures [127, 128, 129] and this mechanism has also been found in a dissipative fiber-laser system [124, 130]. Other mechanisms observed in dissipative systems involve stochastically induced transitions in multistable systems [131] or temporal chaotic dynamics in a non-spatially extended laser with optical injection [132]. Extreme events have been found in a variety of optical cavity systems such as an injected nonlinear optical cavity [133], fiber laser [130, 134], solid-state lasers [135], optical liquid crystal light valve with optical feedback [136], and semiconductor lasers [132, 137, 138]. The role of spatial coupling has not been studied until recently in a pattern-forming optical system composed of a photo-refractive crystal subjected to optical feedback [139, 140] or a low Fresnel number solid-state laser [141]. Extreme events following polarization measurements (vectorial rogue waves) have also been demonstrated in optical fibers [142] and in VCSELs with feedback [143]. Most of these extreme events were characterized from a statistical point of view without establishing their origin from dynamical system theory. Our group demonstrated numerical and experimental results on the formation of extreme events in a spatially extended nonlinear dissipative system and showed that the extreme events occur at the onset of spatiotemporal chaos, hence enabling a link between the statistical and the dynamical properties

of the system.

4.3 Extreme events in a line semiconductor laser

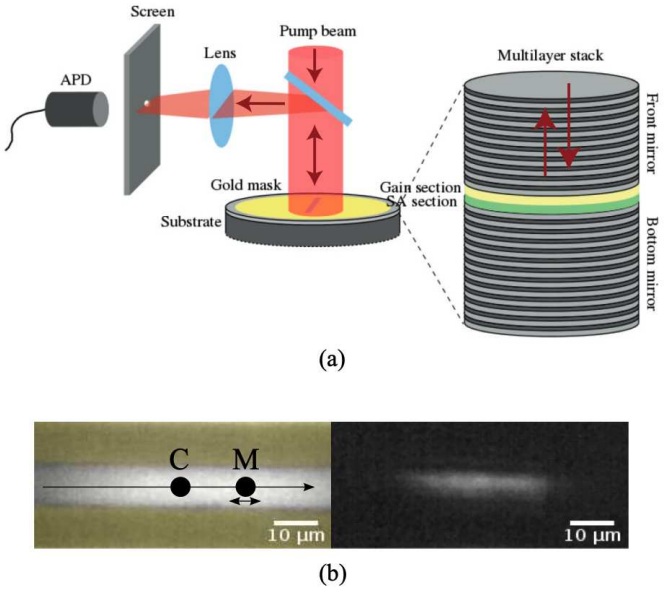


Figure 4.1: Schematic setup and images from experiment to generate and measure extreme events in quasi 1-D system
Panel (a) shows the schematic representation of the experimental setup. Panel (b) shows the top view of the quasi 1-D system below (left) and above (right) threshold. False colour (yellow for the golden mask) has been added to the image below threshold to enhance the contrast of the image. Figure adapted from [144].

Our group (in collaboration with Saliya Coulibaly at the Université de Lille and Marcel G. Clerc at Universidad de Chile) in the past has shown and explained the mechanisms and formation of extreme events in a quasi-1D semiconductor laser [145, 144]. This section of the chapter is inspired by these papers to introduce the readers to the optical system under consideration and characterize the extreme events. The system studied by our group is a planar microcavity laser with an integrated saturable absorber. The device structure is similar to the micropillar lasers outlined earlier on in this thesis but the diameter of the devices under study here is much larger on the order of $\approx$ 100 μm. The planar microcavity is covered with a gold mask and an opening of 80 μm in length and 10μm in width is made to realize the quasi-1D structure. The microcavity is then pumped along this rectangular aperture above threshold and the intensity close to the center and

at a movable point is recorded with a fast avalanche photodiode having a bandwidth of 5 GHz. The temporal signal is then amplified using a low noise high bandwidth amplifier and acquired with a 6 GHz oscilloscope at 20 GS/s or 50 ps sampling time. The movable detector is sequentially displaced to span the entire 1-D structure. The experimental setup and the schematic of the experiment is shown in figure 4.1. In the absence of spatial coupling, this system does not display irregular dynamics and hence, extreme events. Spatial coupling arises from diffraction as a consequence of the large aspect ratio (and Fresnel number) which accommodates a large number of transverse spatial-modes.

The irregular dynamics motivates us to study if such a system is chaotic. The characterization of chaos and spatiotemporal chaos is carried out by studying the Lyapunov exponents or spectrum [146]. These exponents measure the exponential sensitivity to infinitesimally small perturbations to a given trajectory in the phase space. A strictly positive Lyapunov exponent denotes that nearby trajectories will diverge exponentially in time where as a negative exponent signifies that any small perturbation will converge onto a single (original) trajectory. The number of Lyapunov exponents equal the dimensionality of the system under study. A system is said to be chaotic even if a single exponent is strictly positive. Experimentally, we are often in the situation where the underlying dynamical system might be unknown with the desired level of accuracy or we only have access to a limited set of variables describing the dynamical system. It is well known that we can estimate the Lyapunov exponents by using delay coordinate embedding [147, 148]. Some recent results also suggest that we can use reservoir computing [149] to reconstruct the complete dynamics [150]. This technique has also been used to reproduce the dynamics of low and high dimensional dynamical systems and these networks then can be used to infer the full set of Lyapunov exponents [151].

Often, for systems with a known model, the most feasible technique is to model the system under study and numerically estimate the Lyapunov exponents. Let N be the number of discretization points in an extended system described with m degrees of freedom, then the system has $m \times N$ Lyapunov exponents λ_i. If the Lyapunov exponents are sorted in decreasing order and in the thermodynamic limit ($N \to \infty$), these exponents converge to

a continuous spectrum as conjectured by Ruelle [152]. Therefore, if the system displays spatiotemporal chaos in this limit, there exists an infinite number of positive Lyapunov exponents. The set of ordered (descending) Lyapunov exponents provides an upper limit for the strange attractor dimension through the Kaplan-Yorke dimension D_{KY} (equation 4.1) [153], where j is the largest integer that satisfies $\sum_{i=1}^{j} \lambda_i \geq 0$. In the thermodynamic limit, the Kaplan-Yorke dimension diverges with the size of the system as a consequence of the Lyapunov density [154].

$$D_{KY} = j + \sum_{i=1}^{j} \frac{\lambda_i}{\lambda_{j+1}} \tag{4.1}$$

Experimental results aided by numerical simulations demonstrate that the physical origin of extreme events in this system can be attributed to spatiotemporal chaos. The route to spatiotemporal chaos in the system was identified as quasi-periodicity. The fine structure of the largest Lyapunov exponent and of the Kaplan-Yorke dimension D_{KY} helped to infer that intermittency plays an important role in the variation of the number of extreme events found that correlate to the bifurcation of spatiotemporal chaos. Beyond this conclusion, the reader is urged to read the articles from our group cited in the beginning of the section for a detailed discussion on the parameters, establishing spatiotemporal chaos and the numerical schemes used.

In order to proceed, we would first like to introduce the definition we consider for extreme events which is taken from hydrodynamics. In a time trace with many pulses, the height of an event is computed as the change of the relevant variable (intensity here) between two consecutive minima and maxima. For the case of a single pulse embedded in some background variation, one can find a minima ($Mn1$) before the pulse, the maxima ($Mx1$) of the pulse and a subsequent minima ($Mn2$) after the pulse. These three values correspond to one event, the height of the event is then selected using the definition $H = max(|Mn1 - Mx1|, |Mx1 - Mn2|)$. We define a significant height H_s which is the mean of the highest tertile of the probability density function of the height H of the events recorded. Any event having the height greater than twice the significant height is labeled as an extreme event. Another way of expressing this is that any event with an abnormality

index, $AI = H/H_s > 2$ is called an extreme event.

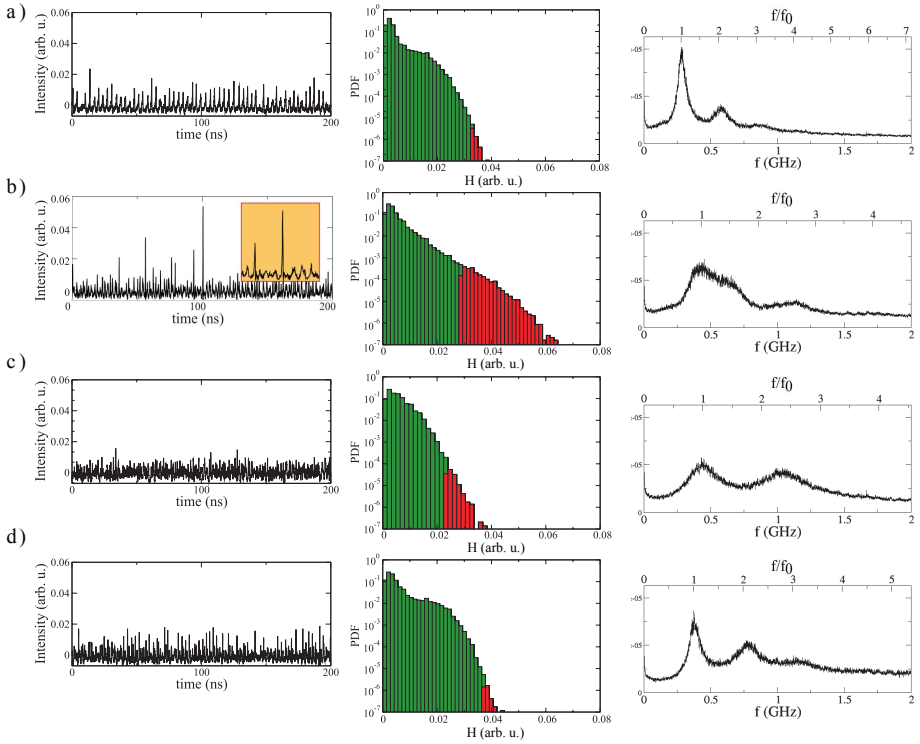


Figure 4.2: Dynamics observed in the quasi-1D system for different values of pump Experimental observation of extreme events, their corresponding logarithm of the PDF of the intensity height H and power spectral density (PSD) for different normalized pump values P/Pth (a) 1.02, (b) 1.17, (c) 1.20, and (d) 1.25 are plotted. Panel (b) also shows an excerpt of the temporal evolution of intensity at the central detector C. Extreme events (AI > 2) are marked in red. Figure adapted from [144]

The results obtained experimentally for different values of pump are shown in figure 4.2. The events labeled in red are the extreme events. The figure also shows the power spectral density and the temporal evolution of the laser intensity to characterize the complex behaviour of the system. For pump values close to the threshold, $P/P_{th} = 1.02$ (subplot a) where P_{th} is the laser threshold, the system exhibits quasiperiodic behaviour and its Fourier spectrum is a well defined peak with a harmonic. In this parameter region, the PDF of the event height decays rapidly and the probability of occurrence of extreme events is very low. With an increase of the pump $P/P_{th} = 1.17$ (subplot b) we observe a large number of extreme events. This can be seen in the non-Gaussian tail of the PDF of the event height. The development of fat tails is also reflected in the Fourier spectrum which shows considerable mixing and broadening of its frequency components. Under

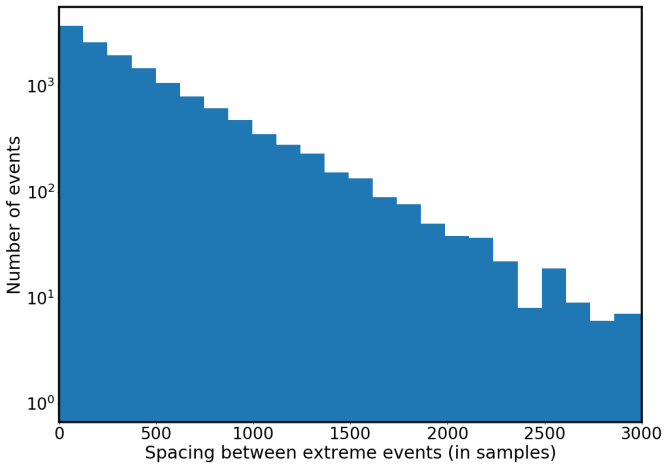


Figure 4.3: Histogram of spacing between two extreme events for $P/P_{th} = 1.2$

further increase of the pump power, $P/P_{th} = 1.2$ (subplot c), intermittent dynamic behaviour in the temporal evolution of the total intensity is replaced by irregular oscillations of the total intensity. The probability of observing an extreme event is reduced and the Fourier spectrum reveals a quasiperiodic behaviour with noise like fluctuations. For further higher values of the pump, $P/P_{th} = 1.25$ (subplot d), the temporal evolution of the intensity shows aperiodic behaviour with the probability of large amplitude events reducing drastically.

The interarrival time for the extreme events is shown in figure 4.3. This figure shows the histogram of time between two consecutive extreme events. It displays Poisson-like statistics characterized by an exponential decrease with mean interspike time $\sigma \approx 21.7\,\mathrm{ns}$ or approximately 435 samples. Hence we can conclude that in the first approximation extreme events are a memoryless Poisson process. The normalized value of pump P/P_{th} in this figure was 1.20. The work on prediction in this chapter will be carried out using the data recorded with the normalize pump of 1.20.

It was shown in ref. [145] that there is a correlation region in space and time in the vicinity of an extreme event. The cross correlation was introduced as follows:

$$X_{c,m}(k) = \frac{1}{N\sigma_{y_c}\sigma_{y_m}} \sum_i [y_c(i) - \bar{y_c}][y_m(i+k) - \bar{y_m}] \tag{4.2}$$

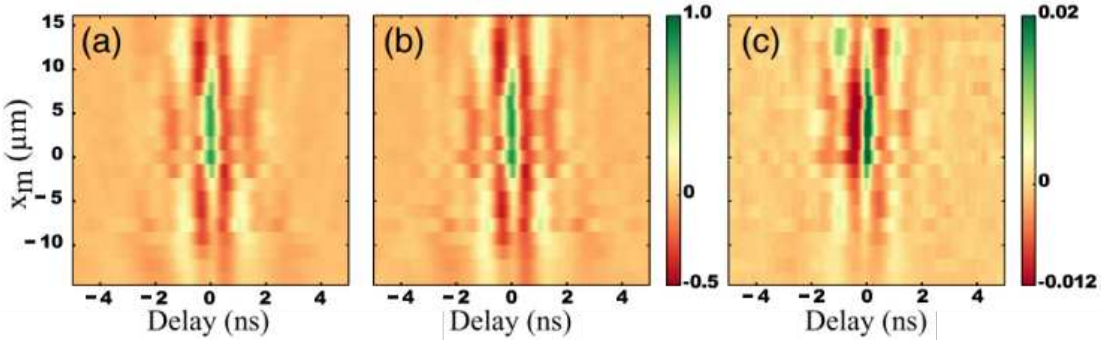


Figure 4.4: Cross-correlation and average plots to identify precursors
(a) Temporal cross-correlation $X_{c,m}(k)$ (defined in text) between detector responses in points C (central detector) and point M at delay k ns. (b) Same as (a) restricted to only extreme events at point C. (c) Average of responses at point M and around the time when an extreme event occurred at C. Adapted from [145]

Where, $X_{c,m}$ is the cross-correlation between the signals recorded by the central detector y_c at point C and the movable detector y_m at point M (in our experiment, $1 \leq m \leq 20$). k is the temporal delay (ns in the plot). The bar symbol and σ indicate the mean value and standard deviation respectively. The results of this are shown in figure 4.4. Subplot (a) was computed for every event at point C. Each event at the central detector is at location (0,0) in this spatiotemporal plot where the y-axis is the span of the quasi-1D system and the x-axis is the time around the event. A high correlation region was identified in space at the time of the extreme event (region in green, centered at time = 0). This region extends to the region smaller than the span of the system and appears vertical with our experimental resolution. This indicates that large amplitude pulse occurs simultaneously in the entire spatially correlated region and disappears rapidly outside this region. In subplot (b), the cross-correlation was computed only only using extreme events at location C and it was observed that the resulting plot is similar to the previous plot. Based on this, it was concluded in ref. [145] that the extreme events are not a result of collision of coherent structures or solitons as this would have resulted in a bending of this correlation region (at least in subplot (b)[1]). Further, it was observed that the correlation bands extend for about 2 ns in time. Subplot (c) plots the average of the responses at point M when an extreme events occurs at C. While this plot is similar in structure to the

[1]This was substantiated through a second plot only limited to extreme events as one could argue that any structure of collision might be washed out in subplot (a) as a result of the disproportionate number of non extreme events. This not the case here.

cross-correlation plot, it shows asymmetry around the correlated structure.

4.4 Prediction and precursors

The prediction of extreme events has wide importance and is prevalent in a variety of complex systems belonging to very diverse fields such as: neuroscience, financial markets, climate, oceanography and geoscience to name a few. Having established that extreme events are a consequence of spatiotemporal chaos in our system, it is thus an interesting endeavour to try and predict the occurrence of these events. The prediction of chaotic dynamics is inherently challenging due to the divergence of nearby trajectories which is captured by the Lyapunov exponents. Thus, in order to have an accurate evolution of the system one would need a very high resolution of measurements[2] which is frequently not the case in experimental data.

In a general case when the system is large, it is not possible to access the observable(s) with arbitrary precision in space and time. It is thus useful to have an estimate about the subspace which contains maximum information about said observable(s). These subspaces or regions will be called precursors and are extremely useful to identify the occurrence of extreme events. Since we consider a 1D structure, the region under consideration is a 2D entity, one dimension for the space and one for time. As a first attempt, we choose to identify precursors using a cross correlation plot presented in the last section. Such a plot might indicate regions from where machine learning algorithms might extract some trends.

The cross-correlation and average plots, while giving us indications of possible precursors, do not identify the direction of flow of information. Moreover, there could exist regions which seem uncorrelated but which would exchange information. To account for such effects, we compute the transfer entropy in collaboration with Saliya Coulibaly at the Université de Lille. The notion of transfer entropy was introduced in ref. [156] as an addition to the concept of mutual information. Mutual information between two systems I and J, M_{IJ} as defined within the paper is the excess amount of code produced by

[2] As an interesting example, one can consider the simulation of the Lorenz system [155] with slightly different numerical precision and the solution will diverge!

erroneously assuming that two systems are independent. While it gives a measure of the information shared between the two systems, by construction the quantity is symmetrical ($M_{IJ} = M_{JI}$) thus contains no directional information. In extended systems it is interesting to identify the rate at which different components exchange information. Transfer entropy answers this question by taking the dynamics of information transport into account thus making it an asymmetrical measure. The original definition can be adapted to our case and the resulting equation is as follows:

$$T_{M\to C}(x_M,\tau) = \sum_{I_{C,n}} p(I_{C,n}, I_{C,n-\tau}, I_{M,n-\tau}) \log_2 \frac{p(I_{C,n}|I_{C,n-\tau}, I_{M,n-\tau})}{p(I_{C,n}|I_{C,n-\tau})} \tag{4.3}$$

Where $T_{M\to C}(x_M,\tau)$ is the transfer of information (in bits) from point M at a certain lag τ to point C. The variable I is the sampled value of the intensity, the first subscript is the location of measurement (*id est* point C for the center and point M for the movable detector) and the second subscript is the time at which the signal is sampled. Therefore, $I_{C,n}$ refers to the intensity sampled at point C at a time t_n. Transfer entropy is plotted in fig. 4.5 and is a function of two coordinates, space (displacement from the center) and time (lags). It indicates the amount of information transfer from that point to the center (which has coordinates (0,0)). In the figure, the x-axis is in micrometers and the y-axis is in sampling time (50 ps or 20 GS/s). Higher values on the logarithmic color bar signify higher information transfer. We observe three main regions of high information transfer. The first region is centered around the center ($x_c - x_m = 0$): this region marks the principle correlated structure around the extreme event. It corresponds to the correlated region (vertical line around $t = 0$) already shown in the correlation plots presented in figure 4.4. The two other regions observed are displaced from the center at about $\pm$ 10 μm from the center. The displaced regions demonstrate a high information transfer at lower lags (further away from the extreme event). This reveals that there is information exchange between regions outside the principle correlation region and the central point (location of the extreme event). While it was weakly visible in the correlation plot, this trend is highlighted clearly here. Additionally, we observe almost no information is transferred from points below -25 lags (or $\approx$ 1.25 ns). This means that we will probably not be able

to extract information about the extreme event occurrence in this region.

This information exchange can be used to identify the region(s) in the 2D space to predict occurrence of extreme events without requiring concurrent measurements throughout the structure. We would like to point out to the readers that the plots from transfer entropy were computed for single points in time, *id est* how intensity recorded at a single point in time at M affects the intensity at C. In principle, we could compute the same quantity for a history of more than one sample, this work is underway.

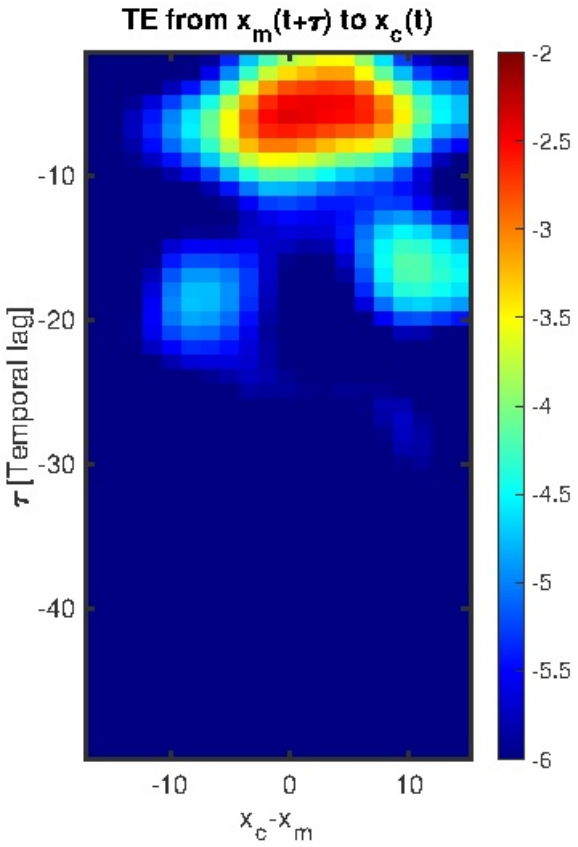


Figure 4.5: Transfer entropy $T_{M\to C}$
Transfer entropy is computed between every point and the center. The x-axis is in micrometers and the y-axis is in samples

Having introduced the different ways of identifying the pertinent information and the direction of flow of information. The next section presents the generation of the dataset used for training different machine learning algorithms. We will limit our analysis for delay times less than 0 (as we want to predict the event before it occurs) and at three points along the structure. These points, P_+, P_0 and P_- are at +12.86 μm from the center, at the center

and at -14.55 μm from the center respectively. We choose these points as P_+ is in a region showing a non zero correlation and a high average while being outside the central correlated structure, P_0 is at the location of the extreme event and P_- is in a region showing almost zero correlation and average. These spatially separated points correspond to regions containing different amounts of information about the extreme event. Figure 4.6 shows the average of the normalized signal recorded at these three points before an event (marked as delay = 0) at the central detector.

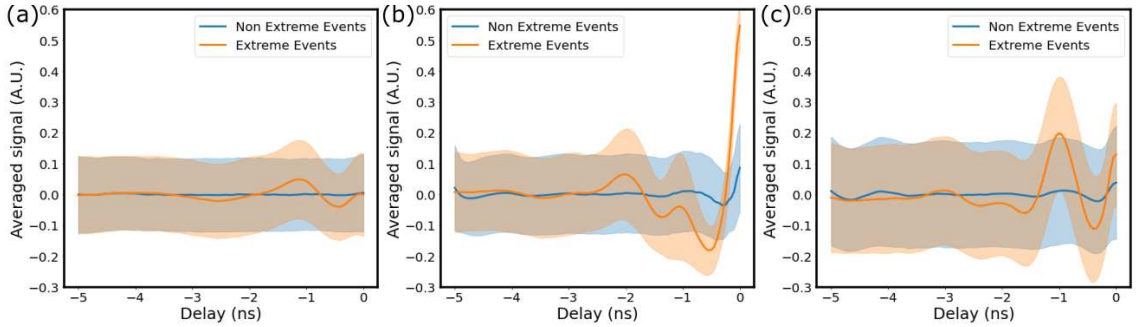


Figure 4.6: Average of signal recorded at three points (a) +12.86 μm from the center, (b) at the center and (c) -14.55 μm from the center.

From the plots presented already, one can intuitively think that predicting using the first two points might result in a reasonable prediction accuracy but the last point should not have any information (or very minimal at best) about the extreme event. In the next section, we show this is indeed the case.

4.5 Prediction with experimental dataset

4.5.1 Dataset generation

In this work, we combine information theory with various machine learning techniques to aid in prediction of extreme events. The experimental dataset used is recorded as mentioned in section 4.2. The total duration of the recording is $\approx$ 2.5 ms. The system is sampled simultaneously at two points, the central point and a moving point. The sampling of the system at 2 points is an experimental constraint as sampling several points simultaneously in the quasi-1D structure can be cumbersome. Ideally simultaneously sampling of the entire system would be desirable but is hardly feasible because it would require a

Streak camera for instance which is not available in the lab (not mentioning the issue of resolution vs recording length and of necessity of single shot recording difficult with low intensities). This sampling is then used to study the information transfer between different locations (movable detector) and the central detector. A dataset is then built by identifying the time τ of all events in the central detector and saving the signal in the movable detector up to 100 samples before said τ. The Lyapunov time in the system is estimated to be on the order of few samples thus the saved history is estimated to be on the order of several tens of Lyapunov times of the system and thus it is a valid assumption that there is no discernible information loss by truncating the history at 100 samples. The events in the central detector are then classified in a binary fashion using the definition of extreme events introduced in the previous section. Since extreme events are rare by definition, we build a balanced dataset by retaining all the extreme events and choosing an equal number of non-extreme events at random. The dataset at the end of the process consists of $2 * N$ time-traces (101 samples each) from the movable detector and a binary label which identifies if the event that occurred at the central detector was an extreme event or not. N is the number of extreme events in the recorded time trace. This same process is repeated for every pair of recorded signals.

We then use various machine learning techniques such as logistic regression, k-nearest neighbours, deep neural networks and reservoir computing to perform cross prediction on the occurrence of an extreme event. These techniques are representative of the different kinds of machine learning techniques used for chaotic time series prediction [6, 157, 158, 159].

The dataset is split in a standard way such that 70% of the data is used as training data and 30% is used as testing data. We can further restrict the information input to the prediction system by implementing a lower and upper bound in time axis *id est* using $I(t)$ for $\tau^{low} \leq t \leq \tau^{up}$. Defining the time running from τ to 0 (and $\tau < 0$) where 0 is the time-stamp of the event in the central detector, lower bound is defined a τ^{low} such that - $\tau < \tau^{low}$ and τ^{up} such that $\tau^{low} < \tau^{up} < 0$. Thus, the time duration of input is m samples in the time between τ_{up} and τ_{low}. The ideal scenario would be when both τ_{up} and m are as low as possible. We then do a sweep in τ_{up} and τ_{low} and present results in a 2-D plot

where every pixel corresponds to classification score[3].

4.5.2 Overview of machine learning techniques used

Logistic Regression

Logistic regression is a popular classification algorithm which uses a logistic function to model a binary dependent variable. Such a classifier can be easily adapted for a multi-class classification using a one v/s all scheme. In such a scheme N parallel classifiers will be trained for N class classification and each of the classifier will determine the probability that a certain example belongs to it's class or not. Thus, as an aggregate, we can achieve a multi-class classification. Since our use case only has two classes, a simple logistic classifier is used. The time series with m samples is fed to the classifier and the predicted output is trained using Eq. 4.4. Where $h_\theta(x)$ is the hypothesis given an input $x \in \mathbb{R}^N$ (column vector) and weights $\theta \in \mathbb{R}^N$ (row vector), y is the label used for training and $\mathrm{Cost(h_\theta(x), y)}$ is the cost function for a given training example. This particular choice of cost function yields a convex optimization problem which minimizes the cost averaged over all the training examples but its derivation is out of scope for the manuscript. The cost function has two contributing parts: when $y = 0$, the cost is $-\log(1 - h_\theta(x))$ and for $y = 1$ the cost is $-\log(h_\theta(x))$. Since $h_\theta(x)$ takes values between 0 and 1, a simple illustration shown in figure 4.7 shows how the cost changes according to the hypothesis $h_\theta(x)$. In panels (a) and (b) the cost is plotted respectively versus hypothesis when $y = 1$ and $y = 0$ and as expected the cost is minimum as the hypothesis approaches the true label. The weights are trained via an iterative procedure using gradient descent. Once the weights are trained, the classifier generates a linear decision boundary such that the output class is 1 if the hypothesis $h_\theta(x)$ for a given x is greater or equal to 1/2. This corresponds to $\theta \cdot x \geq 0$. The output class is determined to be 0 otherwise. The number of trainable parameter are $m + 1$[4].

[3]The term score is a generic one, there can be many ways to define this quantity. The specific definition used will be introduced in the results section of this chapter

[4]m weights + 1 bias. Here $m + 1 = N$, the original time trace of m points is transformed into m+1 points by appending a 1 in front of it, this acts like the bias. An alternative way explicitly mentioning the bias would be to write $h_\theta(x) = \frac{1}{1+e^{(\theta \cdot x)+b}}$

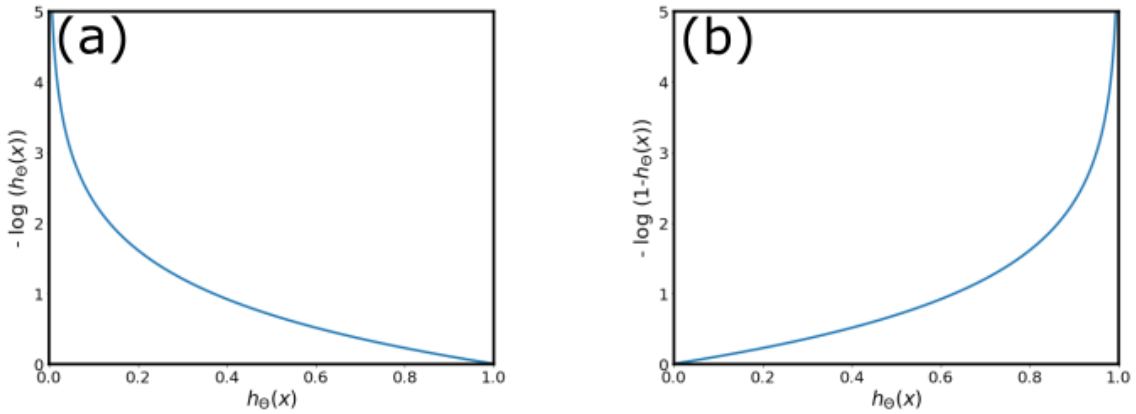


Figure 4.7: Cost per training example
Panels (a) and (b) plot the training cost per example when the y $= 1$ and 0 respectively

$$h_\theta(x) = \frac{1}{1 + e^{\theta \cdot x}}$$
$$\text{Cost}(\text{h}_\theta(\text{x}), \text{y}) = -y \ \log(h_\theta(x)) - (1 - y) \ \log(1 - h_\theta(x)) \tag{4.4}$$

k-Nearest Neighbours

k-Nearest Neighbours (k-NN) is a non parametric algorithm for regression and classification which can be used without making any assumptions about the distribution from which the data is drawn [160]. Previous works have also used this algorithm on vectors generated using delay-coordinate embedding to predict chaotic dynamics [157, 161]. For the binary classification task, the train dataset is first translated into a point cloud in m dimensions via delay embedding and stored in the memory. Every test case also corresponds to one such point in the high dimensional space. The algorithm then computes L^2 norm between every test case and the entire training point cloud and sorts them in ascending order. The top k entries are determined to be the closest neighbours of the test-case and a majority poll among these neighbours determined the class of the test case. This technique is very simple to implement and required no training time. However, the inference time is dependent on the training dataset as the algorithm needs to compute the distances between every test and the entire training dataset. In our use case, we find that $k = 25$ yields optimal results.

Deep Neural Networks

Deep neural networks have become a standard tool for analysis of sequential and static data over the last few years [162, 8, 163]. Out of the several known architectures for deep neural networks, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are particularly interesting for modeling time series data. However, in some machine learning models when the length of time series is not significantly large, feed forward neural networks can also be used as done in ref. [157]. When a feed forward neural network is used for a time series task, the optimizer looks for an optimal function combining inputs from all the time scales available; this doesn't pose a problem if the duration of the time trace is short. However for long time traces, finding relevant information can be very difficult if all the time scales are assumed to have relevant information. It is thus important to relax the constraints and impose fading memory which aids the optimizer to look for physically plausible solutions. Thus for the task at hand, we use a RNN comprising of gated recurrent units (GRUs) which factors in timescales when looking for relevant information. In principle CNNs[5] and feed forward neural networks can also be used for our prediction task as m is on the order of 100 but RNNs were selected as they are most associated with time series analysis.

The schematic of a gated recurrent unit is shown in figure 4.8. I will not introduce the equations behind the GRU but will rather describe the working to aid the reader in comprehending such networks. Each GRU has two gates, *videlicet* update gate $z[t]$ and reset gate $r[t]$. Update gate decides on the relevance of information and helps to capture long-term dependencies in the sequence. The reset gate on the other hand, is used to forget past information. The interplay between the two gates helps capture relevant information at different timescales. At time t, a single GRU cell receives an input $x[t]$ and an input of the hidden state from the previous time step $h[t-1]$. These two are combined and then used to compute the output of the reset gate and update gate via a sigmoid nonlinearity, σ. All these are then used to compute the intermediate hidden state $\hat{h}[t]$. Then via a last operation the final hidden state $h[t]$ is determined to be passed on to the next to the

[5] CNNs have been shown to work well in time series tasks when the time scale of relevant information in known and mentioned in the kernel size.

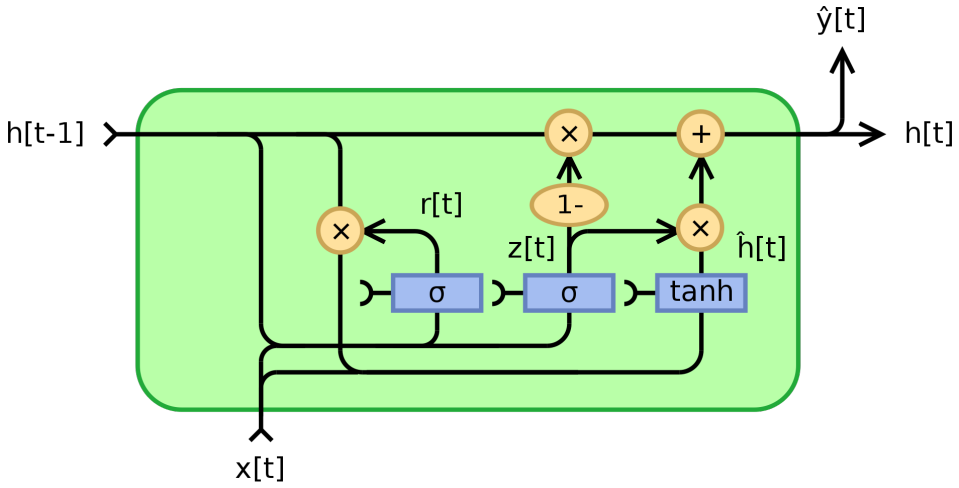


Figure 4.8: Schematic GRU cell

$\hat{h}[t], x[t], r[t], z[t]$ and $h[t]$ are respectively: the intermediate hidden state, input, output from the reset gate, output from the update gate and the hidden state at time t. σ and tanh correspond to the application of the sigmoid and hyperbolic tangent function. Adapted from Wikipedia.

next time step. An optional (final) output $\hat{y}[t]$ is generated when t is less than (equal to) the length of the input. For a single hidden layer and an input dimension of p and output dimension of u, the total number of parameters in the GRU are $3 \times (u^2 + up + u)$. The number of hidden units in an RNN refers to the dimensionality of the hidden state or the dimensionality of the output at any given time. Each hidden layer has hidden cells, as many as the number of time steps and further, each hidden cell is made up of multiple hidden units. For a detailed working of such models, the reader is urged to refer to the original paper introducing GRUs [163].

For the classification task, the appropriate time series is chosen and used to train a 10 instances individual RNNs. Training and predicting with multiple networks with the same dataset minimizes the fluctuation in the prediction score due to random initialization of weights[6]. The architecture of the network is shown in fig 4.9. The default activations of the GRU are retained (sigmoid activation for the reset and update gates and hyperbolic tangent activation for the output of the hidden cell). We use 60 GRU units and the states of these units are read out only at the end of the time trace[7]. The final GRU state is then

[6]The final score on the task is an average of the scores obtained from individual networks. The network parameters are not averaged

[7]Another way of doing it would be to save the states of the GRU at every input time step and this would be the equivalent of recording the transient of the RNN system as we input information. This could be used

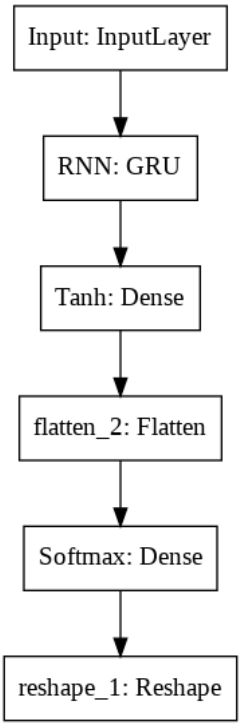


Figure 4.9: Neural network architecture

connected to a dense layer with a hyperbolic tangent activation with 20 *neurons*. The dense layer is then connected to an output layer with softmax activation, $s(z)$[8] (Eq. 4.5) with two neurons which represent the probability of the input belonging to either of the classes.

$$s(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{4.5}$$

The weights are learnt through the backpropagation algorithm using the Adam optimizer and categorical cross entropy as the loss function. The network is trained for 80 epochs and once trained, each test case generates a certain probability for it being an extreme event or not, the class with the highest probability is picked as the final label. We also tried using Leaky-Rectified Linear Unit (LReLU) (Eq. 4.6) as the activation for the dense layer and obtained similar results.

if one wants to predict the changing probability of occurrence of extreme events

[8]In the equation, i denotes the index and K is the number of classes. This output layer transforms real numbers to probability of belonging to a certain class. In our case, $K = 2$.

$$f(x) = \begin{cases} x, & \text{if x} \geq 0 \\ \alpha x, & \text{if x} < 0 \end{cases} \tag{4.6}$$

Reservoir Computing

Reservoir computing has been gaining interest for data driven model discovery and as a consequence model free prediction on various low [6] and high dimensional dynamical systems [151]. They are particularly interesting as reservoirs are themselves dynamical systems with recurrent connections thus making them ideal candidates to map other dynamical systems. A schematic of a reservoir computer is shown in figure 4.10. Additionally the training procedure for a reservoir computer is minimal as compared to its counter parts.

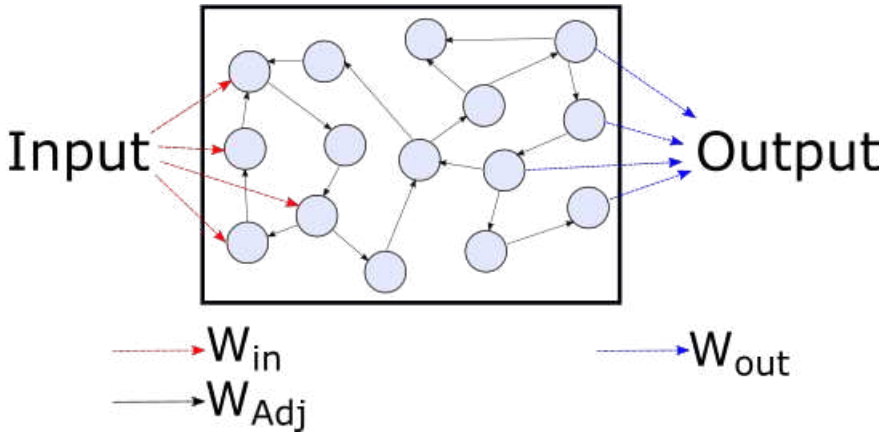


Figure 4.10: Schematic of a reservoir computer
The red and black arrows denote the input weights and the internal reservoir connections respectively. These weights are chosen at random and fixed. The blue arrows correspond to the trainable output weights.

A reservoir consists of three main matrices that control its dynamics. For N nodes in the reservoir, p dimensions of the input vector and m_{op} dimensions of the output, we have input matrix $\mathrm{W_{in}} \in \mathbb{R}^{N \times p}$, the adjacency matrix of the weighted graph, $\mathrm{W_{Adj}} \in \mathbb{R}^{N \times N}$ and the output matrix, $\mathrm{W_{out}} \in \mathbb{R}^{N \times m_{op}}$. Out of these three, only the output matrix is trainable and the rest are chosen at the beginning at random and frozen. Thus the number of trainable parameters are $(N \times m_{op}) + 1$[9]. The reservoir adjacency matrix (binary) is

[9] 1 for bias

generated using a random directed binomial graph for the random connections between the nodes, with an average of 15 connections per node. The weight of each existing edge in the graph is then assigned from a random uniform distribution with the lower and upper bound of $-\mathrm{Res}_{\mathrm{weight}}$ and $\mathrm{Res}_{\mathrm{weight}}$ (weighted adjacency matrix). The value of $\mathrm{Res}_{\mathrm{weight}}$ is determined such that the spectral radius of the adjacency matrix ($W_{Adj} \in R^{N \times N}$) of the reservoir $\approx$ 0.9. The reservoir comprises of 300 nodes each with a hyperbolic tangent activation function. The input weights are drawn from a normal distribution, $\mathcal{N}(0, 0.25)$.

For our classification task, we have l training examples each of length m. We input each training example sequentially (1 sample at a time) into the reservoir using an input matrix W_{in}, the input dimension p is 1 in our case. This results in certain temporal dynamics in the nodes of the reservoir. The states of these nodes are stored at the end of input (after m samples), which in principle should contain information about the temporal order of input samples. Thus, an input time-trace of m samples is converted into a vector of N values. We then train a logistic classifier on the N dimensional vector to learn the weights (W_{out}, m_{op} is also 1 in our case since we predict a binary label) and then the weights are used for inference on the test set.

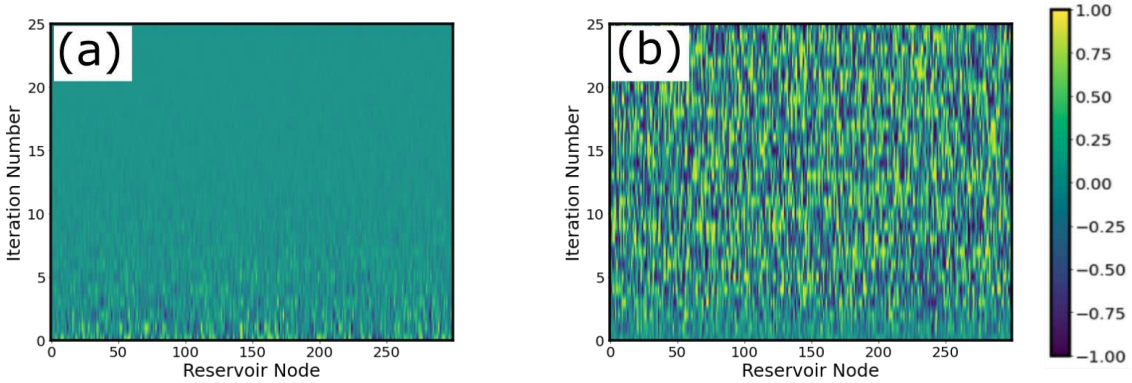


Figure 4.11: Evolution of initial conditions in reservoirs with different spectral radii
The spectral radius in (a) and (b) is approximately 0.9 and 1.8 respectively.

At this point, I would like to introduce a brief interlude on spectral radius and its influence on reservoir computers. As mentioned before, we choose a spectral radius $\approx$ 0.9 for our binary prediction task. While this can seem arbitrary, there is some intuition behind it. The spectral radius is linked to a sort of memory in the system. Considering the reservoir to be a dynamical system, the spectral radius can be tuned to observe a

dynamical system in different regimes. Ideally one would like their reservoir to be on the edge of chaos such that it can satisfy both the approximation and separation property of the system[10]. While this statement is generally true, the optimization of the reservoir like any machine learning algorithm is based on the use case. We further observe that even a chaotic reservoir is capable of reproducing dynamics from certain dynamical systems displaying Dragon-King like extreme events [164]. In the scope of this manuscript, I will limit my analysis to demonstrating the different dynamical regimes of the reservoir used and briefly characterize them. To demonstrate that the value of $\mathrm{Res}_{\mathrm{weight}}$ can act like a bifurcation parameter for the reservoir dynamical system, we choose a certain binary adjacency matrix for the reservoir (as described before) and replace all the nonzero values from a random uniform distribution between -1 and 1 (instead of between $-\mathrm{Res}_{\mathrm{weight}}$ and $\mathrm{Res}_{\mathrm{weight}}$). This directed adjacency matrix ($\mathrm{W}_{\mathrm{Adj}'}$) is then fixed. We can then tune the spectral radius of this matrix by simply multiplying it with a scalar $\mathrm{Res}_{\mathrm{weight}}$[11]. To illustrate the impact of spectral radius, we take two reservoirs with spectral radii $\approx$ 0.9 and 1.8 and start them with an initial condition and observe their dynamics, this is shown in figure 4.11. The evolution equation is as follows (Eq. 4.7):

$$x(n) = F\left(\mathrm{W}_{\mathrm{Adj}} * x(n-1) + \mathrm{W}_{\mathrm{in}} * \delta_{n-1,0}\right) \tag{4.7}$$

Where $x(n)$ is the state of the reservoir (a vector of length N) at the n^{th} iteration and $\delta_{i,j}$ is the Kronecker delta function. Subplot (a) shows how an initial condition decays to a fixed point which is 0 in this case when the spectral radius = 0.9 and subplot(b) shows chaotic dynamics when the spectral radius = 1.8.

Figure 4.12 plots the evolution of one of the reservoir nodes for a better representation of this phenomenon. To further characterize the exponential sensitivity to initial conditions, we perturb each of the reservoirs twice with *nearby* initial conditions and plot the evolution, this divergence in principle must be linked to the largest Lyapunov exponent of the

[10]The approximation and separation property are qualitative metrics linked to a reservoir computer. The approximation property is to emphasize that two *similar* inputs should be mapped to the same output state, thus the reservoir must operate in presence of noise. The separation property suggests that two *sufficiently different* inputs should be mapped to different states. The exact interpretation of these properties is up to the user designing the reservoir computer

[11]$\mathrm{W}_{\mathrm{Adj}} = \mathrm{Res}_{\mathrm{weight}} * \mathrm{W}_{\mathrm{Adj}'}$

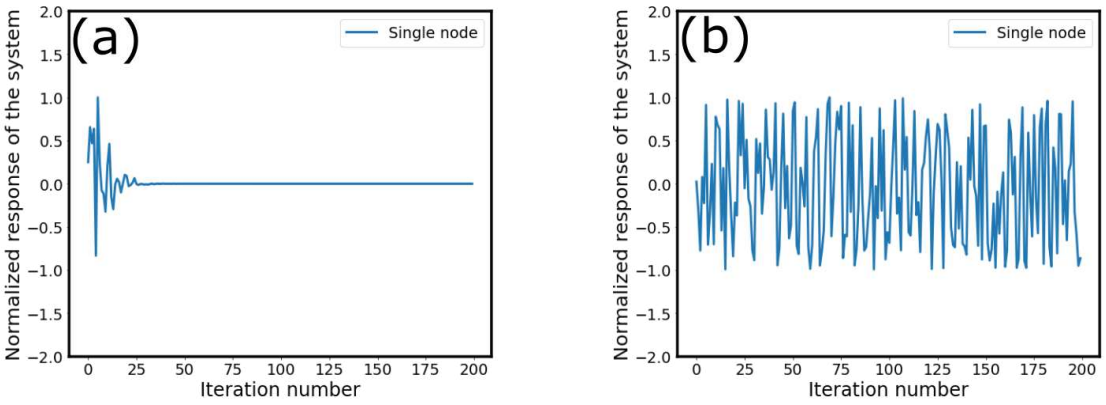


Figure 4.12: Evolution of initial conditions in a single node in reservoirs with different spectral radii

The spectral radius in (a) and (b) is approximately 0.9 and 1.8 respectively.

system. Figure 4.13 illustrates the convergence of the dynamics in subplot (a) (spectral radius $\approx$ 0.9) and exponential divergence in subplot (b) (spectral radius $\approx$ 1.8). Here F^n corresponds to the n^{th} iteration of the reservoir map, x_0 is the initial condition and δ (not to be confused with the Kronecker delta, here δ refers to a small perturbation on the initial condition) is the perturbation to the initial condition. The only intention of these plots is to illustrate how one can obtain different regimes of the reservoir dynamical system using the spectral radius, a rigorous calculation of the Lyapunov exponents of this map are out of scope for this manuscript and would be an independent study in itself. Finally Figure 4.14 shows a 1-D bifurcation plot of one node in the reservoir using the spectral radius as the bifurcation parameter. We can see for a spectral radius less than 1 we observe a fixed point and for a greater spectral radius we observe chaotic dynamics[12].

Based on the discussion, we choose the spectral radius $\approx$ 0.9 such that the reservoir has a fixed point following a transient. We could also choose a spectral radius greater than 1 but then the performance of the reservoir will strongly depend on the length of the input sequence m. This can be intuitively understood from the figure 4.13 (b), the rate of divergence of trajectories is dependent on the spectral radius and thus the approximation property or lack thereof is dependent on the number of iterations of the reservoir update before prediction, since the reservoir is updated for m number of times. Following this

[12]I use the term chaos liberally here. To rigorously prove the observed phenomenon to be spatiotemporal chaos is not the intention

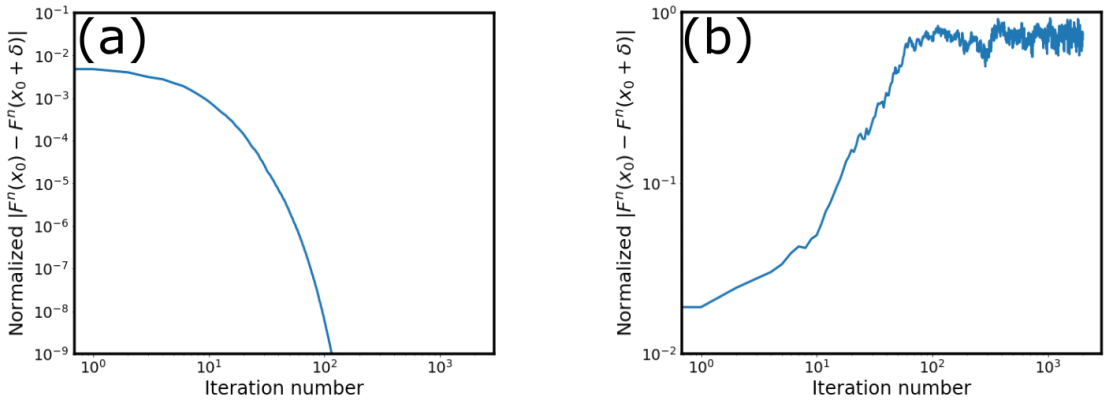


Figure 4.13: Sensitivity to initial conditions in reservoirs with different spectral radii
The spectral radius in (a) and (b) is approximately 0.9 and 1.8 respectively. The difference in trajectories here is averaged over all nodes in the reservoir.

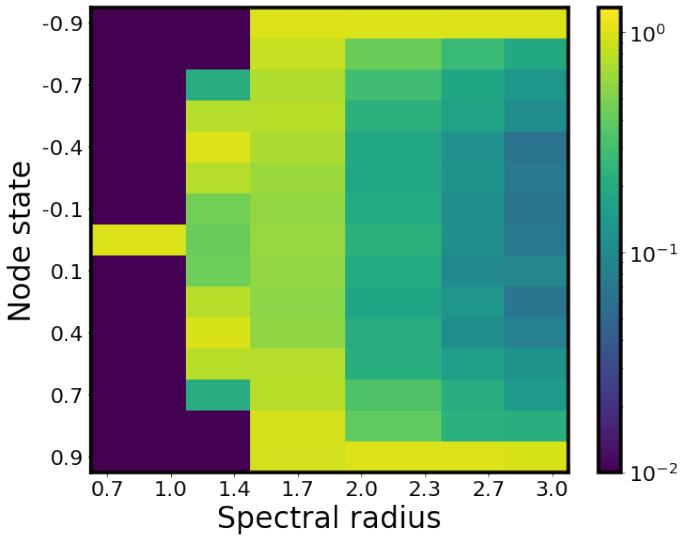


Figure 4.14: 1-D bifurcation diagram of one node in the reservoir using spectral radius as the bifurcation parameter
For generating this plot, the reservoir was initialized at a random state and evolved for 1000 iterations. After this initial transient, the reservoir is assumed to have reached the attractor and them sampled for 200 more iterations to compute the probability of observing it on a particular node state. Computing a similar plot over the entire reservoir yields a similar result but the structure of growth of the attractor size is averaged out. The colourbar represents the probability of observing the node at a given node state.

brief foray into the dynamics of the reservoir and the choice of the spectral radius, the update equation of the reservoir for the binary classification task is as follows (Eq. 4.8):

$$x_l(n) = F\left(W_{Adj} * x_l(n-1) + W_{in} * u_l(n-1)\right) \tag{4.8}$$

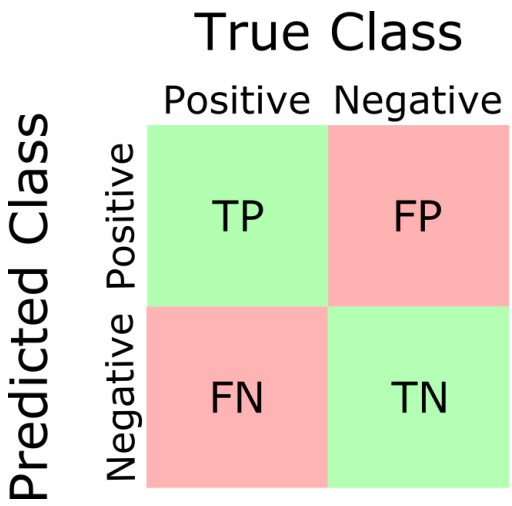


Figure 4.15: Confusion matrix

Where $x_l(u)$ and $u_l(n)$ are the reservoir state vector and input sequence vector at the n^{th} iteration for the l^{th} training example respectively. The initial condition of reservoir dynamical system at the beginning of each new training example, $x(0)$ is a zero valued vector of length N. The training task, once $x_l(m)$ is determined for all training examples, is a logistic regression classification problem as follows:

$$\underset{W_{out}}{\text{Minimize}} \quad \sum_{i=1}^{i=l} \text{Cost}(h_{W_{out}}(x_i(m), y_i) \tag{4.9}$$

Where the functions $\text{Cost}(...)$ and $\text{h}(...)$ have been defined earlier (Eq. 4.4) and y_i corresponds to the binary labels associated for each training task.

4.5.3 Prediction results

Having introduced the different prediction techniques in the previous section, we compare and contrast in this section the results obtained by using them for local and cross prediction tasks. The results of a binary classification task can be expressed in the form of a confusion matrix as shown in figure 4.15. Where TP, TN, FP and FN stand for true positive (extreme event classified as extreme event), true negative (normal event classified as normal event), false positive (normal event classified as extreme event) and false negative (extreme event classified as normal event) respectively. The metric used to judge the classifier is the accuracy defined as $A = \frac{TP+TN}{TP+TN+FP+FN}$. This particular score is

suitable only when the dataset is balanced and a different score needs to be taken for an unbalanced dataset. Considering only a balanced dataset, an ideal classifier will achieve an accuracy of 1 as TP = TN = 1 and FP = FN = 0. A random classifier generates an accuracy of 50 % as all the labels are assigned randomly and TP = TN = FP = FN = 1/2. A biased classifier assigning every event a positive label will achieve an accuracy of 50 % because TP = FP = 1/2 and TN = FN = 0. Similarly for a biased classifier assigning negative labels the accuracy will also be 50 % but with TN = FN =1/2 and TP = FP = 0. The accuracy for the different classification techniques for cross prediction at points P_+, and P_- [13] are shown in figures 4.16 and 4.18 respectively. Prediction at point P_+ in general yields higher accuracy as it is identified to be in the vicinity of the precursor. The maximum achievable accuracy is over 75 % for all the machine learning techniques used. This result is typically obtained for high values of m and τ_{up} as it corresponds to the maximal duration of input trace for prediction in the vicinity of the extreme event. The impact of m and τ_{up} can be individually understood intuitively. Increasing τ_{up} translates to making a prediction closer in time to the extreme event and increasing m increases the input information available to the algorithms to learn and predict extreme events. Consequently, we observe as we increase the time horizon of prediction (decreasing τ_{up}) the prediction accuracy drops and for $\tau_{up} <$ -30 we obtain classification accuracy of $\approx$ 50 %. On the other hand for a fixed τ_{up}, increasing m generally increases the prediction accuracy. A stark example can be seen for lower values of τ_{up}. All the algorithms experience a decrease in prediction accuracy as τ_{up} decreases below -26 and this can be seen as a possible limitation of the temporal horizon for prediction. This value of τ_{up} corresponds to a prediction 1.3 ns before the extreme event.

Prediction at point P_- while having a similar structure as in the previous case (where increasing m and τ_{up} increases accuracy) achieves much lower prediction accuracy. The highest achievable accuracy is in the low 60 percentage points and this quickly drops to approximately 50 % as τ_{up} approaches -11. The poor quality of prediction here is attributed to the fact that this point is calculated to have very little correlation and transfer entropy with the coordinate where the extreme event occurs.

[13] As a recall, points P_+, P_0, and $P_=-$ are at +12.86 μm from the center, at the center and at -14.55 μm from the center respectively

Finally in the case of prediction at point P_0 is shown in figure 4.20. Unlike the previous cases, the maximum possible accuracy achievable is approximately 100 % that is observed for $\tau^{up} = -1$ due to its extreme proximity to the extreme event. As we decrease τ_{up} this accuracy rapidly falls and the local prediction displays comparable performance to the cross prediction at point P_+ with some differences. This figure also displays a structure similar to the previous figures but all the algorithms achieve a much higher prediction accuracy until lower values of τ_{up}. All the algorithms achieve an approximate accuracy of 60% until $\tau_{up} = -31$. Figure 4.22 plots the difference in prediction at point P_+ and P_0. In the region of τ_{up} between -11 and -20 cross prediction outperforms local prediction by about 10 points achieving a maximum accuracy of $\approx$ 65% for low m and achieves similar performance ($\approx$ 70%) for large m [14] For higher and lower values of τ_{up}, the local prediction outperforms cross prediction due to its proximity to extreme events. The better performance of cross prediction can be seen as an interesting case of inferring the evolution of a coupled system with partial measurement: given a limited sampling in space and time, we are able to make useful prediction of the occurrence of extreme events from a location far ($\approx$ 12.86 μm) from it.

[14] Except in the case of Logistic regression. There is no physical explanation behind this difference.

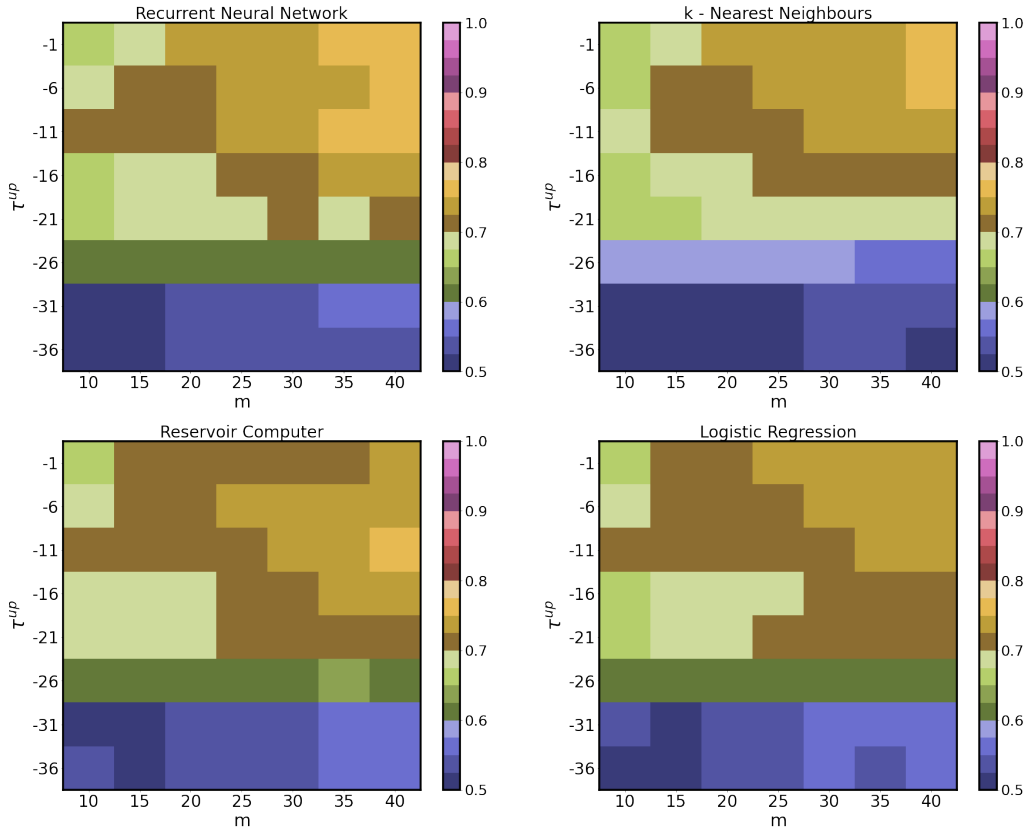


Figure 4.16: Cross prediction accuracy at point P_+ (+12.86 μm from the center)

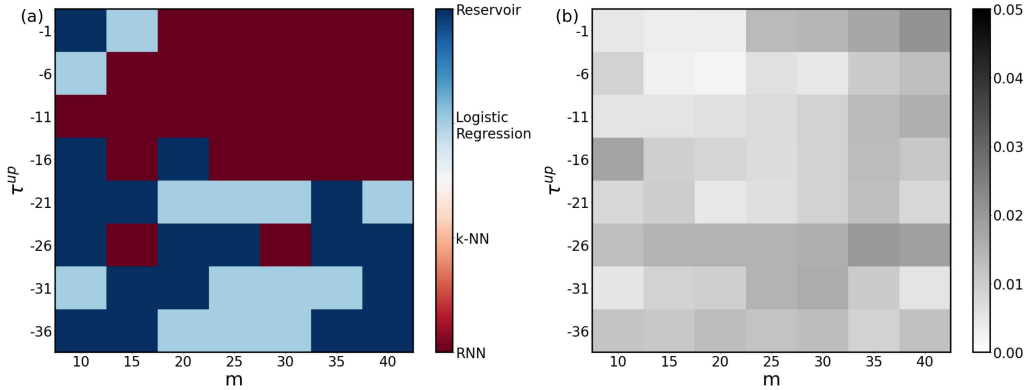


Figure 4.17: Comparing algorithms for cross prediction accuracy at point P_+. Subplots (a) and (b) Indicates which algorithm performs and by how much respectively.

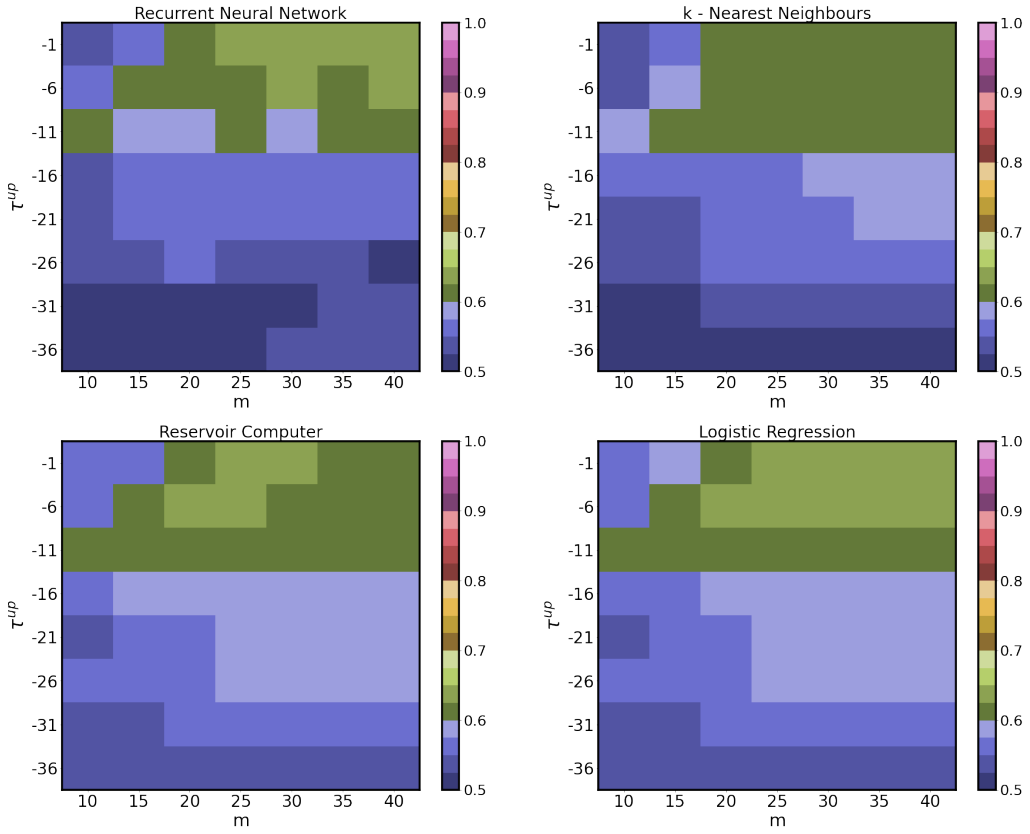


Figure 4.18: Cross prediction accuracy at point P_{-} (-14.55 μm from the center)

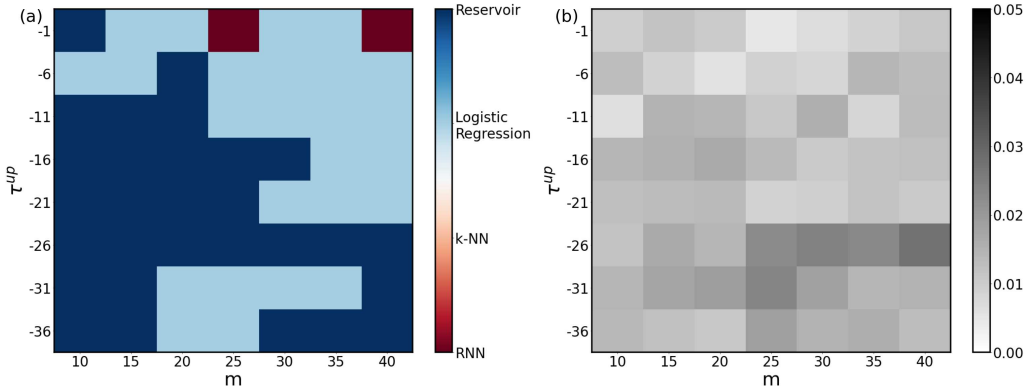


Figure 4.19: Comparing algorithms for cross prediction accuracy at point P_{-}. Subplots (a) and (b) Indicates which algorithm performs and by how much respectively.

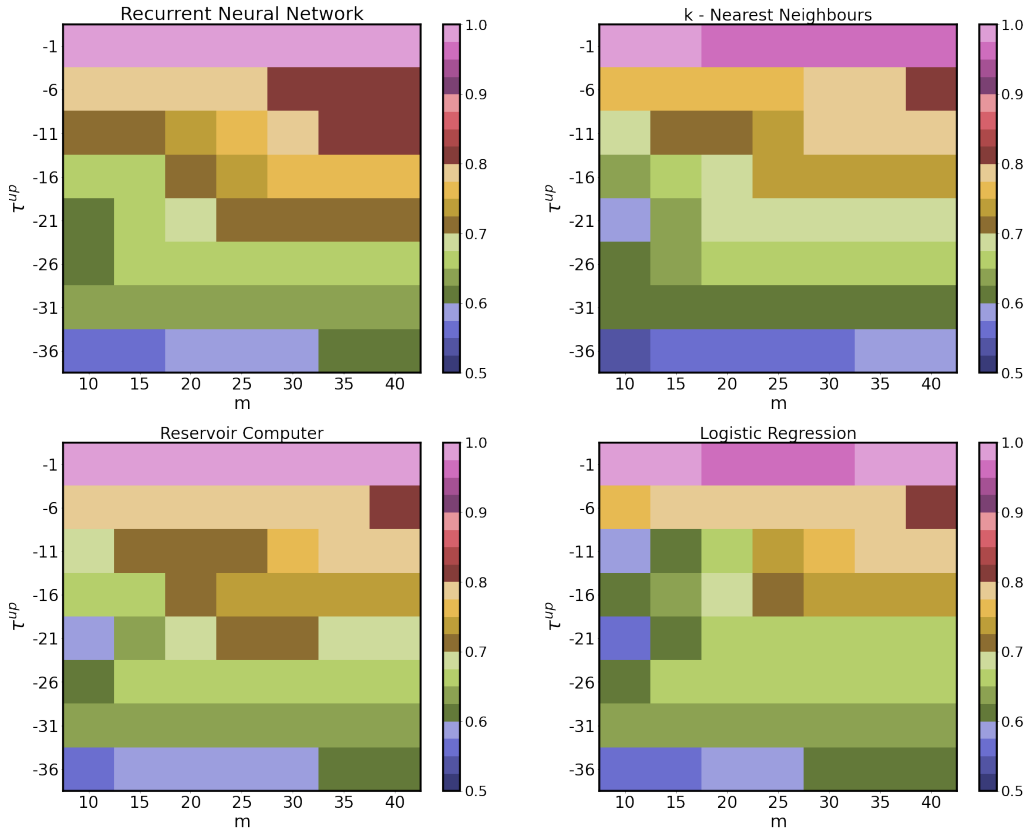


Figure 4.20: Local prediction accuracy

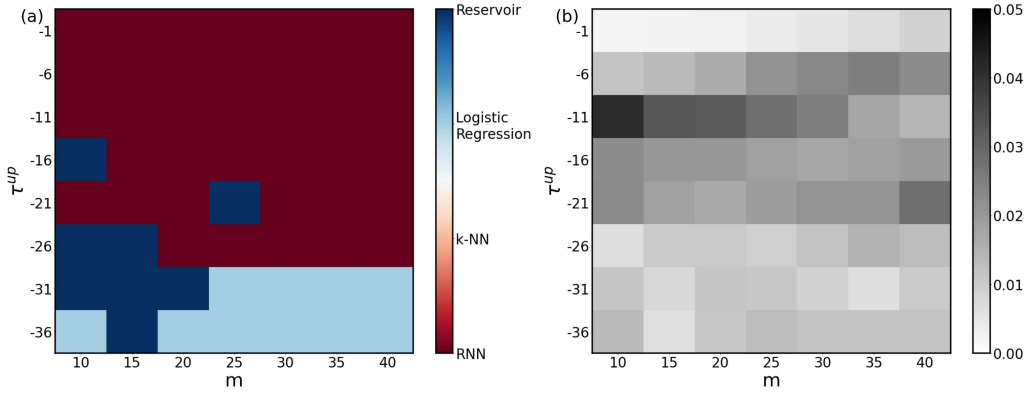


Figure 4.21: Comparing algorithms for local prediction accuracy. Subplots (a) and (b) Indicates which algorithm performs and by how much respectively.

As predicted from the correlation plots and transfer entropy, we observe different locations in space (points P_+, P_0, and P_-) and time (variables m and τ^{up} in each of the accuracy plots) contain different information about the extreme event and this is ultimately reflected in the maximum prediction accuracy achievable. There is a general global structure in m

and τ_{up} that one can observe. While in general increasing m increases information and this intuitively results in higher prediction accuracies, practically the algorithms experience a saturation. This, from a dynamical systems perspective, can be linked to the Lyapunov time and spatiotemporal chaos in the system. A time trace after several Lyapunov times contains no discernible information and can be seen as noise for a prediction algorithm resulting in a saturation in performance, this is what we observe for values of m beyond 40 and hence we stop it there. For much larger values of m, one can expect to see minor decrease in performance efficiency.

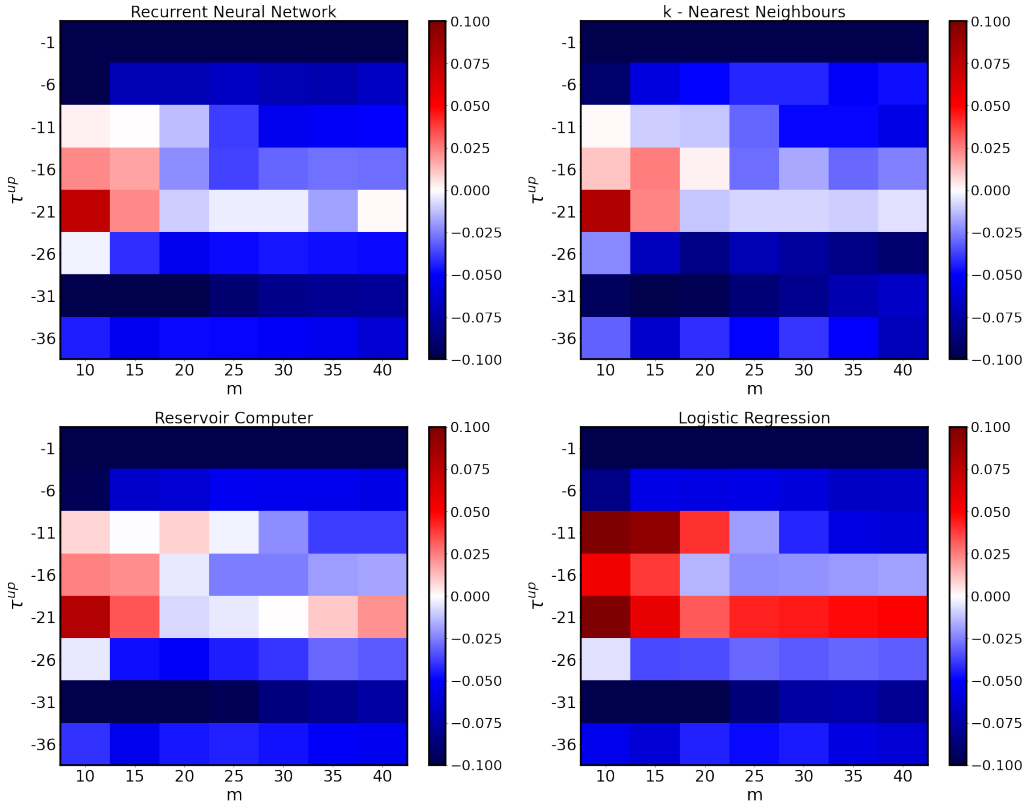


Figure 4.22: A_{Cross} - A_{Local}
Accuracy at point P_+ - Accuracy at point P_0; positive values signify better cross prediction

Further, in order to compare the performance of the different algorithms we generate figures 4.17, 4.19, 4.21 for cross prediction and local prediction respectively. In these figures subplot (a) tracks which algorithm performs best for every parameter combination and subplot (b) plots the difference between the best algorithm and the others[15]. We

[15]This quantity is equal to A_{best} - $\bar{A}'$ where A_{best} is the accuracy of the best algorithm and $\bar{A}'$ is the mean of the accuracies obtained by the other algorithms

would like to point out that while RNNs generally perform slightly better (with the exception of fig. 4.18 where the maximum accuracy is low anyway) they come at the cost of very high number of trainable parameters. The accuracy advantages seem to be in the margin or slightly above the fluctuations in accuracy by altering the datasets, hyperparameters and initial conditions. We further note that the hyperparameters (for reservoir computing and RNNs) were not repeatedly optimized for every parameter combination of m and τ_{up}. Manual hyperparameter tuning can be an extremely time consuming study and was not done extensively during this thesis. Whereas, automatic techniques are very involved and still under development.

Chapter 5

Conclusions

This thesis presented the various research facets surrounding the nonlinear dynamics and computing in the laser stack with saturable absorber first proposed in [35]. The same laser stack can be used to study different dynamical properties based on the coupling type (discrete or continuous) or coupling topology (evanescent or delayed feedback). In this thesis, we demonstrate a variety of experimental and numerical results mainly relating to the computing properties that can be realized using the laser stack. The next section of the chapter will present a chapter wise summary of the key results followed by a section of the possible directions of future work.

5.1 Summary and conclusions

The first chapter presented the motivation for the development of alternative computing primitives and showed how a VCSEL with saturable absorber can be considered to function as an optical neuron due to its neuromorphic properties.

In the second chapter, the effect of coupling via delayed optical feedback was studied. Through experimental and numerical analysis we showed that delayed optical feedback gives rise to attractors in the system and these attractors correspond to different pulsing patterns in the long term depending on the carrier recombination time scales in the gain and saturable absorber regions. These results are in contradiction to the previous work which postulated that such systems with feedback could be used as optical buffers to

store information indefinitely [37, 165]. These attractors can be seen as local minima in a high dimensional space and the system can be seen as performing an energy minimization problem. More specifically, in devices with faster saturable absorber section the final pulsing patterns will be equidistant. Whereas in devices with a faster gain section, the final pulsing patterns can be both equidistant and non-equidistant depending on the feedback time τ of the system. The latter can be seen as an interesting case of occurrence of symmetry broken states in the pulse timings. From an applied point of view, this delay system can be seen as an attractor based computer such as a fully-connected Hopfield network. The initial perturbations that generate the input pulsing pattern play the role of an initial condition and the asymptotic pulsing pattern recorded after several roundtrips in the system indicate which attractor the system has converged to and this convergence is a consequence of similarity between the initial condition and the said asymptotic pulsing pattern. Towards the end of the chapter, we present some results on the polarization dynamics in a single micropillar with and without delayed optical feedback. We showed how a single circular micropillar when subjected to incoherent perturbations can emit pulses with several different polarization states. Further, elliptical micropillars when subjected to similar perturbations emit light along a specific polarization. We then showed the polarization dynamics in a circular micropillar subjected to delayed optical feedback. The delayed optical feedback can be seen as a special case of coherent perturbation. In the presence of delayed optical feedback, a single perturbation from an external source generates a pulse train. We report that in this pulse train the total intensity is constant with noise like fluctuations whereas the linearly polarized intensity of the laser oscillates. The oscillation period appears to be weakly dependent on the bias pump and thus is likely to be of purely dynamical origin.

In the third chapter, we present results on spatial coupling of micropillar lasers. These results are predominantly numerical. We show how altering the pump and/or coupling in 2-D lattices of micropillar lasers can result in interesting circuits such as coincidence detection circuits, temporal pattern recognition circuits, and on chip oscillators. Towards the end, we make a brief foray into computing with neural codes and show how a particular example of neural code called rank order coding can be implemented on using micropillar

lasers. We demonstrate a computing accuracy of $\approx$ 74 % on the MNIST handwritten digit database. While this accuracy is not comparable to state of the art neural networks and other machine learning techniques which report near perfect accuracy, it is noteworthy that the training procedure for rank order coding requires very few samples (20 / class) and is very simple. We strongly believe the accuracy reported here is not a fundamental limit and can be refined with optimal training routines.

The final chapter is devoted to the use of machine learning techniques for extreme event prediction in a system based on the same laser stack. The experimental data is taken from a broad area VCSEL with a gold mask over it realizing a quasi 1-D system. This system, when pumped over threshold, gives rise to extreme events in certain regimes. We developed a method using transfer entropy to identify optimal precursors of extreme events and then trained various machine learning algorithms on it to predict the occurrence of such events ahead in time. An interesting aspect of this is that we could predict the occurrence of extreme events at the center of the quasi 1-D structure using a time trace outside the principle correlated structure of the extreme event. This cross prediction can be seen as a special case of prediction with partial information. We introduced and used several machine learning techniques *videlicet* k-Nearest Neighbours, Logistic Regression, Reservoir Computing, and Recurrent Neural Networks. The maximum prediction accuracy was $\approx$ 75 % for the cross prediction. Our task, since it required the prediction of a binary descriptor (extreme event or non extreme event) using a displaced time series is novel and fundamentally different from other previously reported tasks such as in ref. [158, 159] which predict the amplitude of an extreme event given a certain finely resolved spectral signature.

5.2 Possible directions for future research

From an hardware perspective, there are two main directions for future research. The first to ameliorate the device level performance of the micropillar lasers and the second to scale up computational capacity of micropillar based systems. For the former objective, it would be desirable to develop electrically pumped micropillar lasers as that would

enable easy control of pump over large lattices of micropillars. In order to better harness and master the physical response of the micropillar, it is important to be able to tune the recombination timescales of the carriers in the gain and SA sections. Up to now the parameters were a result of the etching conditions of the micropillars, which result in similar recombination timescales with some fluctuations. Having a much faster SA section would result in larger amplitude excitable pulses without noise-like satellite pulses. Doping of the SA section is a possible way forward to modify the time scale while keeping the other optical properties intact.

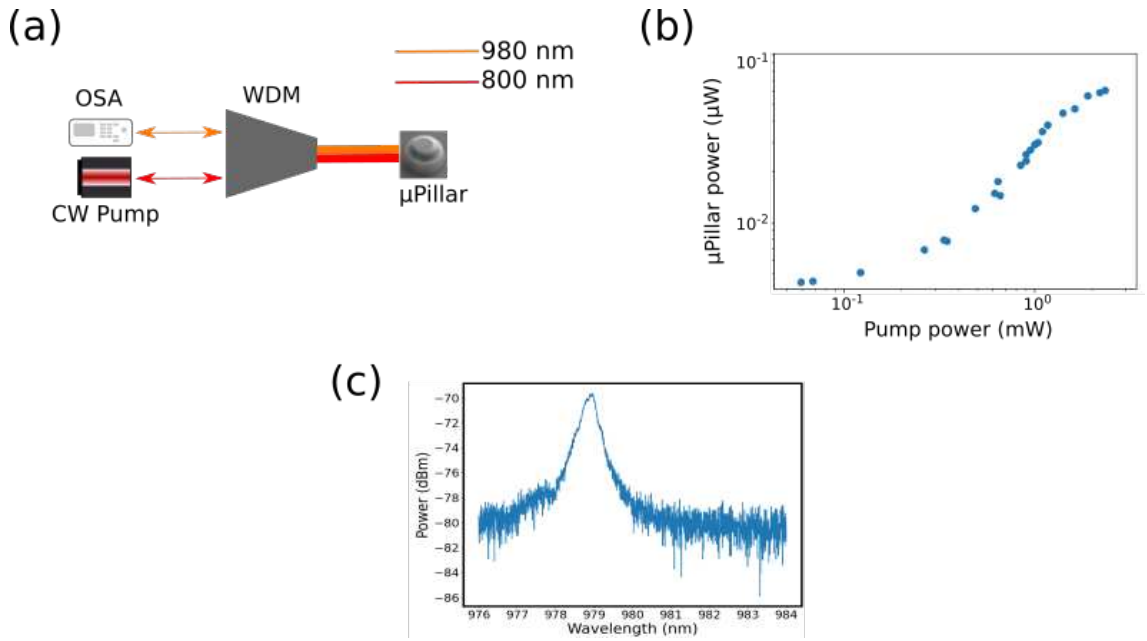


Figure 5.1: Fiber coupled single micropillar laser and preliminary measurement
Subplot (a) shows the optical setup used to characterize the fiber coupled micropillars. OSA: Optical spectrum analyser, WDM: Bidirectional wavelength multiplexer. The bidirectional WDM provides the functionality of merging two fibers carrying different wavelengths of light into one fiber and vice-versa. Subplot (b) plots the laser L-I characteristic on a log log plot. Subplot (c) shows the spectrum of the micropillar laser acquired using an OSA with 20 pm resolution. The 3 dB bandwidth is approximately 0.7 nm.

From a systems perspective, it is imperative to increase the number of coupled micropillar lasers to increase the computational capacity. One of the ways of moving forward would be to couple 2-D lattices using out of plane diffractive optical elements and this would be explored under the framework of the ANR Anaconda project. An alternative way would be to couple individual micropillar lasers to single mode fibers and use the fiber based sys-

tems for increasing the feedback time or coupling several micropillars. We collaborated with an internal startup to develop fiber coupled micropillars and the results are presented in figure 5.1. In subplot (a) we show the schematic of the fiber based setup and in subplot (b) we show the L-I curve. The complete characterization is underway.

On a more fundamental note, we are working towards developing a mathematical model which could explain the modulation observed in the polarization sensitive detection in the delayed optical feedback measurements. The account of the polarization dynamics of the micropillars could be very beneficial since it adds an additional degree of freedom to the system. It allows for properties like neuronal inhibitory dynamics to be optically implemented. The same inhibition dynamics are interesting for computing in networks for spiking neurons but are also necessary for implementing nonlinear decision boundaries such as the XOR logic.

Development in all of the above mentioned aspects would help in making the neuromorphic technology based on micropillar lasers feasible and scalable.

Chapter 6

Synthèse en français

Les travaux présentés dans cette thèse se divisent en deux parties : une première concernant le calcul neuromorphique photonique et une seconde utilisant l'apprentissage automatique appliqué à la prédicition de la dynamique en photonique.

Dans la première partie, nous présentons des résultats obtenus en utilisant des lasers à micropiliers excitables («spiking» ou impulsionnels). Les micropiliers laser excitables ont été démontrés précédemment dans le groupe, grâce à une conception originale de l'empilement des couches du microlaser. La cavité possède une zone active avec deux puits quantiques pour la zone de gain et un puits quantique pour la zone à absorption saturable, et est insérée entre deux miroirs de Bragg. Les trois puits quantiques de la zone active sont en InGaAs/AlGaAs et sont conçus pour une émission et une absorption autour de la résonance de cavité visée à 980nm. Les lasers sont pompés optiquement autour de 800nm, et la zone de gain est pompée sélectivement grâce à une ingénierie spécifique de la structure de l'empilement des miroirs. La compétition entre le gain et les pertes dans le microlaser donne lieu au régime excitable. Les microlasers excitables à absorbant saturable intégré ont des propriétés similaires à celles des neurones biologiques, avec un période réfractaire absolue et relative, un temps de latence à la réponse et la propriété de sommation temporelle, mais fonctionnent à des échelles de temps de l'ordre d'un million de fois plus courtes. Cette similarité nous permet d'appliquer directement certains concepts de calcul biologiquement plausible dans des micropiliers lasers, ouvrant ainsi la voie à de nouveaux substrats pour le calcul neuromorphique ultrarapide

dans le domaine optique. Nous présentons des résultats expérimentaux et numériques sur ces micropiliers, indépendants ou couplés, soit par rétroaction optique retardée, soit par couplage spatial. Dans le cas d'un laser à micropilier avec une rétroaction optique retardée, nous montrons que le système supporte des trains d'impulsions régénératifs qui peuvent être manipulés avec des impulsions optiques uniques. Au temps court, les trains d'impulsions régénératifs sont identiques à la séquence de perturbation initiale et se répètent à la période du retard optique, se comportant ainsi comme un buffer optique. Cependant, dans le long terme, le système présente un degré fini de multistabilité. Toute séquence arbitraire d'impulsions initiales inscrite dans la cavité externe finit par converger vers l'un des régimes stables possibles qui sont constitués d'états périodiques et d'états à symétrie brisée. Les états à symétrie brisée apparaissent comme des régimes périodiques avec des impulsions non équidistantes, brisant ainsi la symétrie de décalage temporel des trains d'impulsions équidistants. Le processus de convergence démontré ici peut être utilisé pour construire des mémoires adressables par leur contenu, comme celles basées sur les réseaux de Hopfield. Nous présentons des résultats expérimentaux sur la dynamique de polarisation de la réponse des lasers micropiliers avec et sans rétroaction. Sans rétroaction optique, le laser à micropiliers peut émettre différents états de polarisation en fonction de sa géométrie. Cependant, avec une rétroaction optique retardée, le micropilier laser émet un train d'impulsions avec une amplitude constante mais un angle de polarisation oscillant, probablement dû à un phénomène d'origine dynamique. Ces résultats ouvrent la voie à l'utilisation du degré de liberté de la polarisation pour le traitement neuromorphique de l'information dans les lasers à micropiliers.

En utilisant des micropiliers couplés spatialement, nous démontrons numériquement divers circuits intégrables sur puce capables d'effectuer des opérations logiques basées sur des impulsions, de reconnaître des séquences temporelles d'impulsions et de générer des trains périodiques. Ces fonctionnalités s'observent soit en modifiant à l'étape de fabrication le couplage spatial entre les micropiliers, soit pendant l'expérience en agissant sur le pompage des micropiliers. Enfin, nous présentons des résultats sur la reconnaissance de chiffres manuscrits en utilisant un codage temporel appelé codage par ordre de rang implémenté sur un ensemble de lasers micropiliers.

Dans la deuxième partie, nous présentons des résultats numériques sur la prédiction de l'occurrence d'événements extrêmes en utilisant des données expérimentales issues d'une microcavité laser de grande extension spatiale. La microcavité utilisée est de conception analogue à celle utilisée dans la première partie de la thèse mais se distingue par le fait que qu'elle est étendue dans une direction, donnant lieu à une structure transverse quasi-1D. Le laser est pompé au dessus de son seuil, et dans certains régimes de paramètres permet l'observation de chaos spatio-temporel et d'événements extrêmes. Notre prédiction est basée sur une connaissance partielle du champ spatio-temporel dans le laser et sur l'identification de précurseurs. Ces derniers ont été identifiés en calculant la corrélation et l'entropie de transfert. Nous avons analysé les performances de plusieurs techniques d'apprentissage automatique telles que la régression logistique, les k-voisins les plus proches, le calcul à réservoir et les réseaux neuronaux récurrents sur la tâche de prédiction. Nous démontrons une bonne capacité de prévision (jusqu'à une précision d'environ 75%) en utilisant les données d'une région spatialement déconnectée de l'emplacement de l'événement extrême.

Appendix A

Publication and conferences

Publication record as of July, 2021.

A.1 Journal articles

- V. A. Pammi, K. Alfaro-Bittner, M. G. Clerc, and S. Barbay, “Photonic Computing With Single and Coupled Spiking Micropillar Lasers” IEEE J. Sel. Topics Quantum Electron., **26**, 1, (2019). DOI: 10.1109/JSTQE.2019.2929187

- S. Terrien*, V. A. Pammi*, N. G. R. Broderick, R. Braive, G. Beaudoin, I. Sagnes, B. Krauskopf, and S. Barbay, “Equalization of pulse timings in an excitable microlaser system with delay” Phys. Rev. Research, **2**, 023012, (2020). DOI: 10.1103/PhysRevResearch.2.023012. *: Authors contributed equally

- S. Terrien, V. A. Pammi, B. Krauskopf, N. G. Broderick, and S. Barbay, “Pulse-timing symmetry breaking in an excitable optical system with delay” Phys. Rev. E, **103**, 012210, (2021) DOI:10.1103/PhysRevE.103.012210

- V. A. Pammi, A. H. Masominia, S. Coulibaly, M. G. Clerc, and S. Barbay, "Precursor-aided prediction of extreme events in a spatio-temporally chaotic system" In prep. (2021)

A.2 Conference submissions

- V. A. Pammi, S. Terrien, N. G. R. Broderick, B. Krauskopf and S. Barbay, "Short and Long Term Memory in Regenerative Spiking Micropillar Lasers" Conference on Lasers and Electro-Optics Europe European Quantum Electronics Conference (CLEO/Europe-EQEC), (2019). DOI:10.1109/CLEOE-EQEC.2019.8873079. (Oral presenter)

- V. A. Pammi, S. Terrien, N. G. R. Broderick, B. Krauskopf, and S. Barbay "Pulse train dynamics in an excitable microcavity laser with delayed optical feedback", The International Symposium on Physics and Applications of Laser Dynamics, (2019). (Oral presenter)

- V. A. Pammi, S. Terrien, N. G. R. Broderick, R. Braive, G. Beaudoin, I. Sagnes, B. Krauskopf, and S. Barbay "Associative memory in regenerative spiking micropillar lasers", Proc. SPIE 11356, Semiconductor Lasers and Laser Dynamics IX, 113560V, (2020). DOI:10.1117/12.2555025. (Oral presenter)

- V. A. Pammi, R. Braive, G. Beaudoin, I. Sagnes, and S. Barbay "Regenerative spike train dynamics in neuromimetic micropillar lasers", Cognitive computing: Merging concepts with hardware, (2018). (Poster presentation)

A.3 Periodicals

- V. A. Pammi and S. Barbay, "Micro-lasers for neuromorphic computing" Photoniques, **104**, 26, (2020). DOI:10.1051/photon/202010426

Bibliography

[1] F. Rosenblatt, “The perceptron: A probabilistic model for information storage and organization in the brain,” *Psychological Review*, vol. 65, no. 6, pp. 386–408, nov 1958.

[2] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[3] A. L. Hodgkin and A. F. Huxley, “Action potentials recorded from inside a nerve fibre [8],” pp. 710–711, 1939.

[4] ——, “A quantitative description of membrane current and its application to conduction and excitation in nerve,” *The Journal of Physiology*, vol. 117, no. 4, pp. 500–544, aug 1952.

[5] W. Maass, T. Natschläger, and H. Markram, “Real-time computing without stable states: A new framework for neural computation based on perturbations,” *Neural Computation*, vol. 14, no. 11, pp. 2531–2560, nov 2002.

[6] H. Jaeger and H. Haas, “Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication,” *Science*, vol. 304, no. 5667, pp. 78–80, apr 2004.

[7] S. J. Cook, T. A. Jarrell, C. A. Brittin, Y. Wang, A. E. Bloniarz, M. A. Yakovlev, K. C. Nguyen, L. T.-H. Tang, E. A. Bayer, J. S. Duerr, H. E. Bülow, O. Hobert, D. H. Hall, and S. W. Emmons, “Whole-animal connectomes of both Caenorhabditis elegans sexes,” *Nature*, vol. 571, no. 7763, pp. 63–71, jul 2019.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[9] T. F. De Lima, H. T. Peng, A. N. Tait, M. A. Nahmias, H. B. Miller, B. J. Shastri, and P. R. Prucnal, "Machine Learning with Neuromorphic Photonics," *Journal of Lightwave Technology*, vol. 37, no. 5, pp. 1515–1534, 2019.

[10] I. Boybat, M. Le Gallo, S. R. Nandakumar, T. Moraitis, T. Parnell, T. Tuma, B. Rajendran, Y. Leblebici, A. Sebastian, and E. Eleftheriou, "Neuromorphic computing with multi-memristive synapses," *Nature Communications*, vol. 9, no. 1, pp. 1–12, jun 2018.

[11] T. Hirtzlin, M. Bocquet, J. O. Klein, E. Nowak, E. Vianello, J. M. Portal, and D. Querlioz, "Outstanding Bit Error Tolerance of Resistive RAM-Based Binarized Neural Networks," in *Proceedings 2019 IEEE International Conference on Artificial Intelligence Circuits and Systems, AICAS 2019*. Institute of Electrical and Electronics Engineers Inc., mar 2019, pp. 288–292.

[12] J. Torrejon, M. Riou, F. A. Araujo, S. Tsunegi, G. Khalsa, D. Querlioz, P. Bortolotti, V. Cros, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M. D. Stiles, and J. Grollier, "Neuromorphic computing with nanoscale spintronic oscillators," *Nature*, vol. 547, no. 7664, pp. 428–431, jul 2017.

[13] D. Marković, A. Mizrahi, D. Querlioz, and J. Grollier, "Physics for neuromorphic computing," *Nature Reviews Physics 2020 2:9*, vol. 2, no. 9, pp. 499–510, jul 2020.

[14] D. A. B. Miller, "Waves, modes, communications, and optics: a tutorial," *Advances in Optics and Photonics, Vol. 11, Issue 3, pp. 679-825*, vol. 11, no. 3, pp. 679–825, sep 2019.

[15] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, "Parallel photonic information processing at gigabyte per second data rates using transient states," *Nature Communications*, vol. 4, no. 1, pp. 1–7, jan 2013.

[16] T. Zhou, X. Lin, J. Wu, Y. Chen, H. Xie, Y. Li, J. Fan, H. Wu, L. Fang, and Q. Dai, "Large-scale neuromorphic optoelectronic computing with a reconfigurable diffractive processing unit," *Nature Photonics*, vol. 15, no. 5, pp. 367–373, may 2021.

[17] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljacic, "Deep learning with coherent nanophotonic circuits," *Nature Photonics*, vol. 11, no. 7, pp. 441–446, jun 2017.

[18] J. Feldmann, N. Youngblood, M. Karpov, H. Gehring, X. Li, M. Stappers, M. Le Gallo, X. Fu, A. Lukashchuk, A. S. Raja, J. Liu, C. D. Wright, A. Sebastian, T. J. Kippenberg, W. H. Pernice, and H. Bhaskaran, "Parallel convolutional processing using an integrated photonic tensor core," *Nature*, vol. 589, no. 7840, pp. 52–58, jan 2021.

[19] M. A. Nahmias, T. F. De Lima, A. N. Tait, H. T. Peng, B. J. Shastri, and P. R. Prucnal, "Photonic Multiply-Accumulate Operations for Neural Networks," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 1, jan 2020.

[20] Z. Cheng, C. Ríos, W. H. Pernice, C. David Wright, and H. Bhaskaran, "On-chip photonic synapse," *Science Advances*, vol. 3, no. 9, p. e1700160, sep 2017.

[21] J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. Pernice, "All-optical spiking neurosynaptic networks with self-learning capabilities," *Nature*, vol. 569, no. 7755, pp. 208–214, may 2019.

[22] K. Vandoorne, W. Dierckx, B. Schrauwen, D. Verstraeten, R. Baets, P. Bienstman, and J. V. Campenhout, "Toward optical signal processing using Photonic Reservoir Computing," *Optics Express*, vol. 16, no. 15, p. 11182, jul 2008.

[23] K. Vandoorne, J. Dambre, D. Verstraeten, B. Schrauwen, and P. Bienstman, "Parallel reservoir computing using optical amplifiers," *IEEE Transactions on Neural Networks*, vol. 22, no. 9, pp. 1469–1481, sep 2011.

[24] F. Duport, B. Schneider, A. Smerieri, M. Haelterman, and S. Massar, "All-optical reservoir computing," *Optics Express*, vol. 20, no. 20, p. 22783, sep 2012.

[25] K. Vandoorne, P. Mechet, T. Van Vaerenbergh, M. Fiers, G. Morthier, D. Verstraeten, B. Schrauwen, J. Dambre, and P. Bienstman, "Experimental demonstration of reservoir computing on a silicon photonics chip," *Nature Communications*, vol. 5, no. 1, pp. 1–6, mar 2014.

[26] L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutierrez, L. Pesquera, C. R. Mirasso, and I. Fischer, "Photonic information processing beyond Turing: an optoelectronic implementation of reservoir computing," *Optics Express*, vol. 20, no. 3, p. 3241, jan 2012.

[27] L. Larger, A. Baylón-Fuentes, R. Martinenghi, V. S. Udaltsov, Y. K. Chembo, and M. Jacquot, "High-speed photonic reservoir computing using a time-delay-based architecture: Million words per second classification," *Physical Review X*, vol. 7, no. 1, 2017.

[28] K. Hicke, M. A. Escalona-Morán, D. Brunner, M. C. Soriano, I. Fischer, and C. R. Mirasso, "Information processing using transient dynamics of semiconductor lasers subject to delayed feedback," *IEEE Journal on Selected Topics in Quantum Electronics*, vol. 19, no. 4, 2013.

[29] J. Vatin, D. Rontani, and M. Sciamanna, "Experimental reservoir computing using VCSEL polarization dynamics," *Optics Express*, vol. 27, no. 13, p. 18579, jun 2019.

[30] D. Brunner and I. Fischer, "Reconfigurable semiconductor laser networks based on diffractive coupling," *Optics Letters*, vol. 40, no. 16, p. 3854, aug 2015.

[31] S. Maktoobi, L. Froehly, L. Andreoli, X. Porte, M. Jacquot, L. Larger, and D. Brunner, "Diffractive coupling for photonic networks: How big can we go?" *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 1, jan 2020.

[32] J. Bueno, S. Maktoobi, L. Froehly, I. Fischer, M. Jacquot, L. Larger, and D. Brunner, "Reinforcement learning in a large-scale photonic recurrent neural network," *Optica*, vol. 5, no. 6, p. 756, jun 2018.

[33] A. Saade, F. Caltagirone, I. Carron, L. Daudet, A. Dremeau, S. Gigan, and F. Krzakala, “Random projections through multiple optical scattering: Approximating Kernels at the speed of light,” in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2016-May. Institute of Electrical and Electronics Engineers Inc., oct 2016, pp. 6215–6219.

[34] M. Giudici, C. Green, G. Giacomelli, U. Nespolo, and J. R. Tredicce, “Andronov bifurcation and excitability in semiconductor lasers with optical feedback,” *Physical Review E*, vol. 55, no. 6, pp. 6414–6418, jun 1997.

[35] S. Barbay, R. Kuszelewicz, and A. M. Yacomotti, “Excitability in a semiconductor laser with saturable absorber,” *Optics Letters*, vol. 36, no. 23, p. 4476, dec 2011.

[36] T. Van Vaerenbergh, M. Fiers, P. Mechet, T. Spuesens, R. Kumar, G. Morthier, B. Schrauwen, J. Dambre, and P. Bienstman, “Cascadable excitability in microrings,” *Optics Express*, vol. 20, no. 18, p. 20292, aug 2012.

[37] B. Garbin, J. Javaloyes, G. Tissoni, and S. Barland, “Topological solitons as addressable phase bits in a driven laser,” *Nat. Commun.*, vol. 6, no. 1, pp. –, jan 2015.

[38] A. Hurtado and J. Javaloyes, “Controllable spiking patterns in long-wavelength vertical cavity surface emitting lasers for neuromorphic photonics systems,” *Applied Physics Letters*, vol. 107, no. 24, p. 241103, dec 2015.

[39] M. A. Nahmias, H.-T. Peng, T. F. de Lima, C. Huang, A. N. Tait, B. J. Shastri, and P. R. Prucnal, “A Laser Spiking Neuron in a Photonic Integrated Circuit,” dec 2020.

[40] D. Brunner, M. C. Soriano, and G. V. der Sande, Eds., *Photonic Reservoir Computing*. De Gruyter, 2019.

[41] B. J. Shastri, A. N. Tait, T. F. de Lima, W. H. P. Pernice, H. Bhaskaran, C. D. Wright, and P. R. Prucnal, “Photonics for artificial intelligence and neuromorphic computing,” *Nature Photonics 2021 15:2*, vol. 15, no. 2, pp. 102–114, jan 2021.

[42] G. Wetzstein, A. Ozcan, S. Gigan, S. Fan, D. Englund, M. Soljačić, C. Denz, D. A. B. Miller, and D. Psaltis, "Inference in artificial intelligence with deep optics and photonics," *Nature*, vol. 588, no. 7836, pp. 39–47, dec 2020.

[43] T. Elsass, K. Gauthron, G. Beaudoin, I. Sagnes, R. Kuszelewicz, and S. Barbay, "Control of cavity solitons and dynamical states in a monolithic vertical cavity laser with saturable absorber," *European Physical Journal D*, vol. 59, no. 1, pp. 91–96, jul 2010.

[44] F. Selmi, "Réponse excitable et propriétés neuromimétiques de micropiliers lasers à absorbant saturable," Ph.D. dissertation, Université Paris-Sud XI, 2015.

[45] J. L. A. Dubbeldam, B. Krauskopf, and D. Lenstra, "Excitability and coherence resonance in lasers with saturable absorber," *Phys. Rev. E*, vol. 60, no. 6, pp. 6580–6588, 1999.

[46] M. A. Nahmias, B. J. Shastri, A. N. Tait, and P. R. Prucnal, "A leaky integrate-and-fire laser neuron for ultrafast cognitive computing," *IEEE J. Select. Topics Quantum Electron.*, vol. 19, no. 5, pp. 1–12, 2013.

[47] M. Bache, F. Prati, G. Tissoni, R. Kheradmand, L. A. Lugiato, I. Protsenko, and M. Brambilla, "Cavity soliton laser based on VCSEL with saturable absorber," *Applied Physics B: Lasers and Optics*, vol. 81, no. 7, pp. 913–920, nov 2005.

[48] F. Selmi, R. Braive, G. Beaudoin, I. Sagnes, R. Kuszelewicz, and S. Barbay, "Relative refractory period in an excitable semiconductor laser," *Physical Review Letters*, vol. 112, no. 18, p. 183902, may 2014.

[49] F. Selmi, R. Braive, G. Beaudoin, I. Sagnes, R. Kuszelewicz, T. Erneux, and S. Barbay, "Spike latency and response properties of an excitable micropillar laser," *Physical Review E*, vol. 94, no. 4, p. 042219, oct 2016.

[50] T. Erneux and S. Barbay, "Two distinct excitable responses for a laser with a saturable absorber," *Phys. Rev. E*, vol. 97, no. 6, jun 2018.

[51] F. Selmi, R. Braive, G. Beaudoin, I. Sagnes, R. Kuszelewicz, and S. Barbay, "Temporal summation in a neuromimetic micropillar laser," *Optics Letters*, vol. 40, no. 23, pp. 1–5, dec 2015.

[52] S. Terrien, B. Krauskopf, N. G. R. Broderick, L. Jaurigue, and K. Lüdge, "Q-switched pulsing lasers subject to delayed feedback: A model comparison," *Physical Review A*, vol. 98, no. 4, p. 043819, oct 2018.

[53] E. M. Izhikevich, *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. The MIT press, 2007.

[54] S. Terrien, B. Krauskopf, N. G. R. Broderick, L. Andréoli, F. Selmi, R. Braive, G. Beaudoin, I. Sagnes, and S. Barbay, "Asymmetric noise sensitivity of pulse trains in an excitable microlaser with delayed optical feedback," *Physical Review A*, vol. 96, no. 4, oct 2017.

[55] F. T. Arecchi, G. Giacomelli, A. Lapucci, and R. Meucci, "Two-dimensional representation of a delayed dynamical system," *Physical Review A*, vol. 45, no. 7, p. R4225, apr 1992.

[56] G. Giacomelli and A. Politi, "Multiple scale analysis of delayed dynamical systems," *Physica D: Nonlinear Phenomena*, vol. 117, no. 1-4, pp. 26–42, jun 1998.

[57] F. Marino and G. Giacomelli, "Pseudo-spatial coherence resonance in an excitable laser with long delayed feedback," *Chaos*, vol. 27, no. 11, p. 114302, nov 2017.

[58] B. Kelleher, C. Bonatto, P. Skoda, S. P. Hegarty, and G. Huyet, "Excitation regeneration in delay-coupled oscillators," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 81, no. 3, p. 036204, mar 2010.

[59] B. Romeira, R. Avo, J. M. L. Figueiredo, S. Barland, J. Javaloyes, R. Avó, J. M. L. Figueiredo, S. Barland, J. Javaloyes, R. Avo, J. M. L. Figueiredo, S. Barland, and J. Javaloyes, "Regenerative memory in time-delayed neuromorphic photonic resonators," *Scientific Reports*, vol. 6, no. 1, p. 19510, may 2016.

[60] M. Marconi, J. Javaloyes, S. Balle, and M. Giudici, “How lasing localized structures evolve out of passive mode locking,” *Physical Review Letters*, vol. 112, no. 22, p. 223901, jun 2014.

[61] J. K. Jang, M. Erkintalo, S. Coen, and S. G. Murdoch, “Temporal tweezing of light through the trapping and manipulation of temporal cavity solitons,” *Nature Communications*, vol. 6, no. 1, pp. 1–7, jun 2015.

[62] P. Camelin, J. Javaloyes, M. Marconi, and M. Giudici, “Electrical addressing and temporal tweezing of localized pulses in passively-mode-locked semiconductor lasers,” *Physical Review A*, vol. 94, no. 6, p. 063854, dec 2016.

[63] J. Javaloyes, P. Camelin, M. Marconi, and M. Giudici, “Dynamics of Localized Structures in Systems with Broken Parity Symmetry,” *Physical Review Letters*, vol. 116, no. 13, p. 133901, mar 2016.

[64] S. Terrien, B. Krauskopf, N. G. R. Broderick, R. Braive, G. Beaudoin, I. Sagnes, and S. Barbay, “Pulse train interaction and control in a microcavity laser with delayed optical feedback,” *Optics Letters*, vol. 43, no. 13, pp. 3013–3016, jul 2018.

[65] B. J. Shastri, M. A. Nahmias, A. N. Tait, A. W. Rodriguez, B. Wu, and P. R. Prucnal, “Spike processing with a graphene excitable laser,” *Scientific Reports*, vol. 6, no. 1, p. 19126, may 2016.

[66] J. Foss, A. Longtin, B. Mensour, and J. Milton, “Multistability and Delayed Recurrent Loops,” *Phys. Rev. Lett.*, vol. 76, no. 4, pp. 708–711, jan 1996.

[67] B. W. Connors, “Neuroscience: Single-neuron mnemonics,” *Nature*, vol. 420, no. 6912, pp. 133–134, nov 2002.

[68] R. Chaudhuri and I. Fiete, “Computational principles of memory,” *Nature Neuroscience*, vol. 19, no. 3, pp. 394–403, feb 2016.

[69] B. Krauskopf and J. J. Walker, “Bifurcation Study of a Semiconductor Laser with Saturable Absorber and Delayed Optical Feedback,” in *Nonlinear Laser Dynamics:*

From Quantum Dots to Cryptography, K. Lüdge, Ed. Weinheim, Germany: Wiley-VCH Verlag GmbH Co. KGaA, dec 2012, pp. 161–181.

[70] S. Terrien, B. Krauskopf, and N. G. Broderick, "Bifurcation analysis of the yamada model for a pulsing semiconductor laser with saturable absorber and delayed optical feedback," *SIAM Journal on Applied Dynamical Systems*, vol. 16, no. 2, pp. 771–801, apr 2017.

[71] K. Ikeda and O. Akimoto, "Instability Leading to Periodic and Chaotic Self-Pulsations in a Bistable Optical Cavity," *Phys. Rev. Lett.*, vol. 48, no. 9, pp. 617–620, mar 1982.

[72] S. Yanchuk and P. Perlikowski, "Delay and periodicity," *Phys. Rev. E*, vol. 79, no. 4, p. 46221, 2009.

[73] J. L. A. Dubbeldam and B. Krauskopf, "Self-pulsations of lasers with saturable absorber : dynamics and bifurcations," *Opt. Commun.*, vol. 159, no. 4-6, p. 325, 1999.

[74] K. Engelborghs, T. Luzyanina, and G. Samaey, "DDE-BIFTOOL: a Matlab package for bifurcation analysis of delay differential equations," Department of Computer Science, KU Leuven, Leuven, Belgium, Tech. Rep., 2001.

[75] J. Sieber, K. Engelborghs, T. Luzyanina, G. Samaey, and D. Roose, "DDE-BIFTOOL Manual - Bifurcation analysis of delay differential equations," 2014.

[76] B. Balachandran, T. Kalmár-Nagy, and D. E. Gilsinn, *Delay differential equations: Recent advances and new directions*. Springer US, 2009.

[77] B. Krauskopf and K. Green, "Computing unstable manifolds of periodic orbits in delay differential equations," *J. Comput. Phys.*, vol. 186, pp. 230–249, 2003.

[78] A. Keane and B. Krauskopf, "Chenciner bubbles and torus break-up in a periodically forced delay differential equation," *Nonlinearity*, vol. 31, no. 6, pp. R165–R187, may 2018.

[79] S. Terrien, V. A. Pammi, B. Krauskopf, N. G. Broderick, and S. Barbay, “Pulse-timing symmetry breaking in an excitable optical system with delay,” *Physical Review E*, vol. 103, no. 1, p. 012210, jan 2021.

[80] S. Terrien, V. A. Pammi, N. G. R. Broderick, R. Braive, G. Beaudoin, I. Sagnes, B. Krauskopf, and S. Barbay, “Equalization of pulse timings in an excitable micro-laser system with delay,” *Phys. Rev. Research*, vol. 2, no. 2, 2020.

[81] S. Yanchuk, S. Ruschel, J. Sieber, and M. Wolfrum, “Temporal Dissipative Solitons in Time-Delay Feedback Systems,” *Phys. Rev. Lett.*, vol. 123, no. 5, 2019.

[82] M. Marconi, J. Javaloyes, S. Barland, S. Balle, and M. Giudici, “Vectorial dissipative solitons in vertical-cavity surface-emitting lasers with delays,” *Nature Photonics*, vol. 9, no. 7, pp. 450–455, jun 2015.

[83] A. Hurtado, I. D. Henning, and M. J. Adams, “Optical neuron using polarisation switching in a 1550nm-VCSEL,” *Optics Express*, vol. 18, no. 24, p. 25170, nov 2010.

[84] J. Vatin, D. Rontani, and M. Sciamanna, “Enhanced performance of a reservoir computer using polarization dynamics in VCSELs,” *Optics Letters*, vol. 43, no. 18, p. 4497, sep 2018.

[85] ——, “Experimental realization of dual task processing with a photonic reservoir computer,” *APL Photonics*, vol. 5, no. 8, p. 086105, aug 2020.

[86] S. Xiang, Z. Ren, Y. Zhang, Z. Song, and Y. Hao, “All-optical neuromorphic XOR operation with inhibitory dynamics of a single photonic spiking neuron based on a VCSEL-SA,” *Optics Letters*, vol. 45, no. 5, p. 1104, mar 2020.

[87] V. Klinshov, L. Lücken, D. Shchapin, V. Nekorkin, and S. Yanchuk, “Multistable Jittering in Oscillators with Pulsatile Delayed Feedback,” *Phys. Rev. Lett.*, vol. 114, no. 17, p. 178103, apr 2015.

[88] ——, “Emergence and combinatorial accumulation of jittering regimes in spiking oscillators with delayed feedback,” *Phys. Rev. E*, vol. 92, no. 4, p. 042914, oct 2015.

[89] C. V. Vreeswijk, L. F. Abbott, G. B. Ermentrout, C. Van Vreeswijk, L. F. Abbott, and G. Bard Ermentrout, “When inhibition not excitation synchronizes neural firing,” *J. Comput. Neurosci.*, vol. 1, no. 4, pp. 313–321, dec 1994.

[90] P. C. Bressloff and S. Coombes, “A Dynamical Theory of Spike Train Transitions in Networks of Integrate-and-Fire Oscillators,” *SIAM J. Appl. Math.*, vol. 60, no. 3, pp. 820–841, jan 2000.

[91] P. Grelu and J. M. Soto-crespo, *Temporal soliton "molecules" in mode-locked lasers: Collisions, pulsations, and vibrations*. Springer, Berlin, Heidelberg, 2008, vol. 751.

[92] K. Krupa, K. Nithyanandan, U. Andral, P. Tchofo-Dinda, and P. Grelu, “Real-Time Observation of Internal Motion within Ultrafast Dissipative Optical Soliton Molecules,” *Phys. Rev. Lett.*, vol. 118, no. 24, 2017.

[93] B. Romeira, J. M. L. Figueiredo, and J. Javaloyes, “Delay dynamics of neuromorphic optoelectronic nanoscale resonators: Perspectives and applications,” *Chaos*, vol. 27, no. 11, p. 114323, nov 2017.

[94] H.-T. Peng, T. F. de Lima, M. A. Nahmias, A. N. Tait, B. J. Shastri, and P. R. Prucnal, “Autaptic Circuits of Integrated Laser Neurons,” in *2019 Conference on Lasers and Electro-Optics (CLEO)*. Washington, D.C.: OSA, 2019, pp. 1–2.

[95] G. V. der Sande, D. Brunner, and M. C. Soriano, “Advances in photonic reservoir computing,” *Nanophotonics*, vol. 6, no. 3, pp. 561–576, 2017.

[96] J. Robertson, E. Wade, Y. Kopp, J. Bueno, and A. Hurtado, “Toward Neuromorphic Photonic Networks of Ultrafast Spiking Laser Neurons,” *IEEE J. Sel. Topics Quantum Electron.*, vol. 26, no. 1, pp. 1–15, 2020.

[97] J. J. Hopfield, “Neural networks and physical systems with emergent collective computational abilities.” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 79, no. 8, pp. 2554–2558, apr 1982.

[98] T. Heuser, M. Pflüger, I. Fischer, J. A. Lott, D. Brunner, and S. Reitzenstein, “Developing a photonic hardware platform for brain-inspired computing based on 5×5 VCSEL arrays,” *JPhys Photonics*, vol. 2, no. 4, p. 44002, aug 2020.

[99] T. Van Vaerenbergh, K. Alexander, J. Dambre, and P. Bienstman, “Excitation transfer between optically injected microdisk lasers,” *Optics Express*, vol. 21, no. 23, p. 28922, nov 2013.

[100] S. Barbay, I. Sagnes, R. Kuszelewicz, A. M. Yacomotti, and R. Braive, “Localized states and excitability in a monolithic VCSEL with saturable absorber,” *2011 5th Rio De La Plata Workshop on Laser Dynamics and Nonlinear Photonics, LDNP 2011*, 2011.

[101] K. Alfaro-Bittner, S. Barbay, and M. G. Clerc, “Pulse propagation in a 1D array of excitable semiconductor lasers,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 30, no. 8, p. 083136, aug 2020.

[102] V. A. Pammi, K. Alfaro-Bittner, M. G. Clerc, and S. Barbay, “Photonic Computing With Single and Coupled Spiking Micropillar Lasers,” *IEEE J. Sel. Topics Quantum Electron.*, vol. 26, no. 1, pp. 1–7, jan 2019.

[103] W. W. Chow, S. W. Koch, and M. Sargent, *Semiconductor-Laser Physics*. Springer Berlin Heidelberg, 1994.

[104] H. T. Peng, G. Angelatos, T. F. de Lima, M. A. Nahmias, A. N. Tait, S. Abbaslou, B. J. Shastri, and P. R. Prucnal, “Temporal information processing with an integrated laser neuron,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 1, jan 2020.

[105] A. Rodan and P. Tiňo, “Minimum complexity echo state network,” *IEEE Transactions on Neural Networks*, vol. 22, no. 1, pp. 131–144, jan 2011.

[106] G. M. Shepherd and R. K. Brayton, “Logic operations are properties of computer-simulated interactions between excitable dendritic spines,” *Neuroscience*, vol. 21, no. 1, pp. 151–165, 1987.

[107] Á. Tóth and K. Showalter, “Logic gates in excitable media,” *The Journal of Chemical Physics*, vol. 103, no. 6, pp. 2058–2066, 1995.

[108] R. Toth, C. Stone, B. d. L. Costello, A. Adamatzky, and L. Bull, “Simple Collision-Based Chemical Logic Gates with Adaptive Computing,” in *Theoretical and Technological Advancements in Nanotechnology and Molecular Computation*. IGI Global, may 2011, pp. 162–175.

[109] A. Jacobo, D. Gomilla, P. Colet, and M. Matias, “Logical operations using excitable cavity solitons,” in *CLEO/Europe and EQEC 2009 Conference Digest*, 2009.

[110] A. Jacobo, D. Gomila, M. A. Matías, and P. Colet, “Logical operations with localized structures,” *New Journal of Physics*, vol. 14, jan 2012.

[111] R. W. Gerchberg and W. O. Saxton, “A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures,” *OPTIK*, vol. 35, no. 2, pp. 237–246, 1972.

[112] S. Thorpe, D. Fize, and C. Marlot, “Speed of processing in the human visual system,” *Nature*, vol. 381, no. 6582, pp. 520–522, 1996.

[113] S. Thorpe and J. Gautrais, “Rank Order Coding,” in *Computational Neuroscience*. Springer US, 1998, pp. 113–118.

[114] G. Nicolis and C. Nicolis, *Foundations of Complex Systems: Emergence, Information and Prediction*, 2nd ed. NJ: World Scientific, 2012.

[115] L. Draper, “"Freak" Ocean Waves,” *Oceanus (X:4)*, jul 1964.

[116] S. Haver, “Evidences of the Existence of Freak Waves,” in *Rogue Waves*, Brest, France, 2000.

[117] R. Höhmann, U. Kuhl, H. J. Stöckmann, L. Kaplan, and E. J. Heller, “Freak waves in the linear regime: A microwave study,” *Physical Review Letters*, vol. 104, no. 9, p. 093901, mar 2010.

[118] V. B. Efimov, A. N. Ganshin, G. V. Kolmakov, P. V. McClintock, and L. P. Mezhov-Deglin, “Rogue waves in superfluid helium,” *European Physical Journal: Special Topics*, vol. 185, no. 1, pp. 181–193, 2010.

[119] K. Dysthe, H. Socquet-Juglard, K. Trulsen, H. E. Krogstad, and J. Liu, “"Freak" waves and large-scale simulations of surface gravity waves,” in *“Freak” waves and large-scale simulations of surface gravity waves*, Honolulu, 2005.

[120] P. C. Liu and K. R. Machutchon, “Are there different kinds of rogue waves?” in *OMAE2006, 25th Int. Conf. Offshore Mechanics and Arctic Engineering*, Hamburg, jun 2006, pp. 1–6.

[121] A. Chabchoub, N. Hoffmann, M. Onorato, and N. Akhmediev, “Super roguewaves: Observation of a higher-order breather inwaterwaves,” *Physical Review X*, vol. 2, no. 1, p. 011015, mar 2012.

[122] D. R. Solli, C. Ropers, P. Koonath, and B. Jalali, “Optical rogue waves,” *Nature*, vol. 450, no. 7172, pp. 1054–1057, dec 2007.

[123] J. M. Dudley, G. Genty, and B. J. Eggleton, “Harnessing and control of optical rogue waves in supercontinuum generation,” *Optics Express*, vol. 16, no. 6, p. 3644, mar 2008.

[124] A. Mussot, A. Kudlinski, M. Kolobov, E. Louvergneaux, M. Douay, and M. Taki, “Observation of extreme temporal events in CW-pumped supercontinuum,” *Optics Express*, vol. 17, no. 19, p. 17010, sep 2009.

[125] F. T. Arecchi, U. Bortolozzo, A. Montina, and S. Residori, “Granularity and inhomogeneity are the joint generators of optical rogue waves,” *Physical Review Letters*, vol. 106, no. 15, p. 153901, apr 2011.

[126] B. Kibler, J. Fatome, C. Finot, G. Millot, F. Dias, G. Genty, N. Akhmediev, and J. M. Dudley, “The Peregrine soliton in nonlinear fibre optics,” *Nature Physics*, vol. 6, no. 10, pp. 790–795, aug 2010.

[127] A. Antikainen, M. Erkintalo, M. Dudley, G. Genty, A. Antikainen, M. Erkintalo, J. M. Dudley, and G. Genty, “On the phase-dependent manifestation of optical rogue waves,” *Nonlinearity*, vol. 25, no. 7, p. R73, jul 2012.

[128] S. Birkholz, E. T. Nibbering, C. Brée, S. Skupin, A. Demircan, G. Genty, and G. Steinmeyer, “Spatiotemporal rogue events in optical multiple filamentation,” *Physical Review Letters*, vol. 111, no. 24, p. 243903, dec 2013.

[129] D. Pierangeli, F. Di Mei, C. Conti, A. J. Agranat, and E. DelRe, “Spatial Rogue Waves in Photorefractive Ferroelectrics,” *Physical Review Letters*, vol. 115, no. 9, p. 093901, aug 2015.

[130] C. Lecaplain, P. Grelu, J. M. Soto-Crespo, and N. Akhmediev, “Dissipative rogue waves generated by chaotic pulse bunching in a mode-locked laser,” *Physical Review Letters*, vol. 108, no. 23, p. 233901, jun 2012.

[131] A. N. Pisarchik, R. Jaimes-Reátegui, R. Sevilla-Escoboza, G. Huerta-Cuellar, and M. Taki, “Rogue waves in a multistable system,” *Physical Review Letters*, vol. 107, no. 27, p. 274101, dec 2011.

[132] C. Bonatto, M. Feyereisen, S. Barland, M. Giudici, C. Masoller, J. R. Leite, and J. R. Tredicce, “Deterministic optical rogue waves,” *Physical Review Letters*, vol. 107, no. 5, p. 053901, jul 2011.

[133] A. Montina, U. Bortolozzo, S. Residori, and F. T. Arecchi, “Non-gaussian statistics and extreme waves in a nonlinear optical cavity,” *Physical Review Letters*, vol. 103, no. 17, p. 173901, oct 2009.

[134] S. Randoux and P. Suret, “Experimental evidence of extreme value statistics in Raman fiber lasers,” *Optics Letters*, vol. 37, no. 4, p. 500, feb 2012.

[135] M. G. Kovalsky, A. A. Hnilo, and J. R. Tredicce, “Extreme events in the Ti:sapphire laser,” *Optics Letters*, vol. 36, no. 22, p. 4449, nov 2011.

[136] M. G. Clerc, G. González-Cortés, and M. Wilson, “Extreme events induced by spatiotemporal chaos in experimental optical patterns,” *Optics Letters*, vol. 41, no. 12, p. 2711, jun 2016.

[137] A. Karsaklian Dal Bosco, D. Wolfersberger, and M. Sciamanna, “Extreme events in time-delayed nonlinear optics,” *Optics Letters*, vol. 38, no. 5, p. 703, mar 2013.

[138] É. Mercier, A. Even, E. Mirisola, D. Wolfersberger, and M. Sciamanna, “Numerical study of extreme events in a laser diode with phase-conjugate optical feedback,” *Physical Review E*, vol. 91, no. 4, p. 042914, apr 2015.

[139] V. Odent, M. Taki, and E. Louvergneaux, “Experimental spatial rogue patterns in an optical feedback system,” *Natural Hazards and Earth System Science*, vol. 10, no. 12, pp. 2727–2732, 2010.

[140] N. Marsal, V. Caullet, D. Wolfersberger, and M. Sciamanna, “Spatial rogue waves in a photorefractive pattern-forming system,” *Optics Letters*, vol. 39, no. 12, p. 3690, jun 2014.

[141] C. Bonazzola, A. Hnilo, M. Kovalsky, and J. R. Tredicce, “Optical rogue waves in an all-solid-state laser with a saturable absorber: Importance of the spatial effects,” *Journal of Optics (United Kingdom)*, vol. 15, no. 6, p. 064004, jun 2013.

[142] F. Baronio, B. Frisquet, S. Chen, G. Millot, S. Wabnitz, and B. Kibler, “Observation of a group of dark rogue waves in a telecommunication optical fiber,” *Physical Review A*, vol. 97, no. 1, p. 013852, jan 2018.

[143] C.-H. Uy, D. Rontani, and M. Sciamanna, “Vectorial extreme events in VCSEL polarization dynamics,” *Optics Letters*, vol. 42, no. 11, p. 2177, jun 2017.

[144] S. Coulibaly, M. G. Clerc, F. Selmi, and S. Barbay, “Extreme events following bifurcation to spatiotemporal chaos in a spatially extended microcavity laser,” *Physical Review A*, vol. 95, 2017.

[145] F. Selmi, S. Coulibaly, Z. Loghmari, I. Sagnes, G. Beaudoin, M. G. Clerc, and S. Barbay, "Spatiotemporal Chaos Induces Extreme Events in an Extended Microcavity Laser," *Physical Review Letters*, vol. 116, no. 1, p. 013901, jan 2016.

[146] P. Manneville, "Dissipative structures and weak turbulence," in *Chaos — The Interplay Between Stochastic and Deterministic Behaviour*. Springer Berlin Heidelberg, apr 1995, pp. 257–272.

[147] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. Cambridge, England: Cambridge University Press, 2003.

[148] U. Parlitz and C. Merkwirth, "Prediction of Spatiotemporal Time Series Based on Reconstructed Local States," *Physical Review Letters*, vol. 84, no. 9, p. 1890, feb 2000.

[149] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Computer Science Review*, vol. 3, no. 3, pp. 127–149, aug 2009.

[150] Z. Lu, J. Pathak, B. Hunt, M. Girvan, R. Brockett, and E. Ott, "Reservoir observers: Model-free inference of unmeasured variables in chaotic systems," *Chaos*, vol. 27, no. 4, p. 041102, apr 2017.

[151] J. Pathak, Z. Lu, B. R. Hunt, M. Girvan, and E. Ott, "Using machine learning to replicate chaotic attractors and calculate Lyapunov exponents from data," *Chaos*, vol. 27, no. 12, 2017.

[152] D. Ruelle, "Large volume limit of the distribution of characteristic exponents in turbulence," *Communications in Mathematical Physics*, vol. 87, no. 2, pp. 287–302, jun 1982.

[153] E. Ott, *Chaos in Dynamical Systems*. Cambridge University Press, aug 2002.

[154] M. R. Paul, M. I. Einarsson, P. F. Fischer, and M. C. Cross, "Extensive chaos in Rayleigh-Bénard convection," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 75, no. 4, p. 045203, apr 2007.

[155] E. N. Lorenz, “Deterministic Nonperiodic Flow,” *Journal of Atmospheric Sciences*, vol. 20, no. 2, pp. 130–148, 1963.

[156] T. Schreiber, “Measuring information transfer,” *Physical Review Letters*, vol. 85, no. 2, pp. 461–464, jul 2000.

[157] P. Amil, M. C. Soriano, and C. Masoller, “Machine learning algorithms for predicting the amplitude of chaotic laser pulses,” *Chaos*, vol. 29, no. 11, p. 113111, 2019.

[158] M. Närhi, L. Salmela, J. Toivonen, C. Billet, J. M. Dudley, and G. Genty, “Machine learning analysis of extreme events in optical fibre modulation instability,” *Nature Communications 2018 9:1*, vol. 9, no. 1, pp. 1–11, nov 2018.

[159] L. Salmela, C. Lapre, J. M. Dudley, and G. Genty, “Machine learning analysis of rogue solitons in supercontinuum generation,” *Scientific Reports 2020 10:1*, vol. 10, no. 1, pp. 1–8, jun 2020.

[160] N. S. Altman, “An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression,” *The American Statistician*, vol. 46, no. 3, p. 175, aug 1992.

[161] J. D. Farmer and J. J. Sidorowich, “Predicting chaotic time series,” *Physical Review Letters*, vol. 59, no. 8, pp. 845–848, aug 1987.

[162] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, nov 1997.

[163] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, “Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling,” dec 2014.

[164] H. L. Hugo, M. Oriá, D. Sornette, E. Ott, and D. J. Gauthier, “Predictability and suppression of extreme events in a chaotic system,” *Physical Review Letters*, vol. 111, no. 19, p. 198701, nov 2013.

[165] B. Romeira, R. Avo, J. M. Figueiredo, S. Barland, and J. Javaloyes, “Regenerative memory in time-delayed neuromorphic photonic resonators,” *Scientific Reports*, vol. 6, jan 2016.

université PARIS-SACLAY | ÉCOLE DOCTORALE Ondes et matière (EDOM)

Titre : Calcul photonique avec des micropiliers excitables couplés et prédiction d'événements extrêmes dans des microcavités lasers

Mots clés : Calcul neuromorphique, Laser excitable, Micropilier laser, Cavités couplées, Dynamique à retard, Prédiction d´événements extrêmes

Résumé : Dans la première partie de la thèse, nous présentons des résultats expérimentaux et numériques sur le calcul neuromorphique photonique obtenus avec des micropiliers lasers excitables (impulsionnels ou "spiking") couplés. La dynamique des lasers excitables présente plusieurs similitudes avec celle des neurones biologiques, tout en possédant des échelles de temps beaucoup plus rapides. Cela nous permet d'appliquer certains concepts de calcul inspirés des neurosciences directement aux micropiliers lasers, ouvrant ainsi la voie à de nouveaux substrats neuromorphiques en photonique. Les micropiliers lasers peuvent être couplés soit par rétroaction optique retardée, soit spatialement. Dans le cas d'une rétroaction optique retardée, nous montrons que le système supporte des trains régénératifs d'impulsions qui peuvent être manipulés (inscrits, effacés, déplacés) avec des perturbations optiques uniques. Au temps court, le système se comporte comme une mémoire tampon optique préservant l'information temporelle sur la séquence de perturbations initiales. Cependant, nous démontrons que dans le long terme, toute séquence d'impulsions arbitraires inscrite doit converger vers l'un des états multistables accessibles au système. Les états multistables sont des trains d'impulsions réguliers et périodiques ou à symétrie temporelle brisée, en fonction des paramètres physiques du micropilier laser. Les états à symétrie temporelle brisée apparaissent comme des séquences d'impulsions non régulières qui peuvent être contrôlées expérimentalement. Le comportement asymptotique des états multistables peut être utilisé pour construire des mémoires adressables par leur contenu, comme celles basées sur les réseaux de Hopfield. Nous présentons également des résultats expérimentaux sur la dynamique de polarisation des impulsions excitables émises par les micropiliers lasers avec et sans rétroaction, ouvrant la voie à l'utilisation du degré de liberté de polarisation pour le traitement neuromorphique. Dans le cas des micropiliers couplés spatialement, nous introduisons et étudions numériquement des circuits intégrables sur puce capables d'effectuer des opérations en logique excitable, de reconnaître des séquences temporelles d'impulsions et de générer des trains d'impulsions périodiques. Enfin, nous présentons un algorithme de reconnaissance de caractères utilisant un codage basé sur l'ordre temporel d'émission des impulsions dans un ensemble de micropiliers lasers. La deuxième partie de la thèse est consacrée à la prédiction numérique de l'occurrence d'événements extrêmes à l'aide de techniques d'apprentissage automatique et utilisant des données expérimentales obtenues dans un laser à semiconducteurs en régime de chaos spatio-temporel. La prédiction se base sur une connaissance partielle du champ spatio-temporel dans le laser et sur l'identification de précurseurs. Les précurseurs ont été identifiés grâce à l'entropie de transfert. En utilisant des techniques d'apprentissage automatique, nous avons pu prévoir les événements extrêmes avec une grande précision en utilisant des données de précurseurs provenant d'une région spatialement déconnectée du lieu de l'événement extrême. L'horizon de prédiction ainsi obtenu va au-delà des échelles de corrélations spatiale et temporelle présentes dans les données initiales.

université PARIS-SACLAY | ÉCOLE DOCTORALE Ondes et matière (EDOM)

Title: Photonic computing with coupled spiking micropillars and extreme event prediction in microcavity lasers

Keywords: Neuromorphic computing, Spiking lasers, Micropillar lasers, Coupled cavities, Delay dynamics, Extreme events forecasting

Abstract: The work presented in this thesis can be divided into two parts: photonic neuromorphic computing and machine learning applied to photonics. In the first part of the thesis, we present experimental and numerical results developed using coupled excitable (spiking) micropillar lasers. Spiking lasers exhibit several similarities to biological neurons but operate at much faster timescales. This similarity enables us to directly map some biologically plausible computing frameworks onto micropillar lasers thus paving the way for new neuromorphic substrates in the optical domain. The micropillar lasers can be coupled either using delayed optical feedback or spatial coupling. We first consider a single micropillar laser with delayed optical feedback. In this case, we show that the system is capable of supporting regenerative spike trains which can be manipulated (written, erased, and displaced) with single optical perturbations. In the short term, the system behaves like an optical buffer preserving the timing information about the initial perturbation sequence. However, we demonstrate that in the long term any arbitrary perturbation pattern eventually converges to one of the possible multistable spiking patterns. This finite degree of multistability consists of periodic and symmetry-broken states depending on the physical parameters of the micropillar laser. Symmetry-broken states appear as pulsing patterns with non-regular timings which can be controlled with experimental parameters. The process of convergence together with the multistability demonstrated here can be used to build content-addressable memories such as the ones based on Hopfield networks. We also present experimental results on the polarization dynamics of the response from micropillar lasers with and without feedback opening the way to using the polarization degree of freedom for neuromorphic processing. Using spatially coupled micropillars, we demonstrate numerically various circuits integrable on-chip capable of performing spike-based logic operations, spiking pattern recognition, and generating periodic spike trains. Lastly, we present results on handwritten digit recognition using a spike coding scheme called rank order coding implemented on an ensemble of micropillar lasers. The second part of the thesis is devoted to the numerical prediction of the occurrence of extreme events by using experimentally recorded data from a quasi 1-D semiconductor laser displaying spatio-temporal chaos. Our prediction is based on partial information of the spatio-temporal field in the laser system and on the identification of precursors. The precursors were identified using correlation plots and transfer entropy. Using machine learning techniques, we demonstrate extreme events forecasting with high accuracy using precursor data from a region spatially disconnected from the location of the extreme event. The prediction horizon thus obtained goes beyond the spatial and temporal correlation scales present in the initial data.