Université Paris Cité - Politecnico di Torino

Image classification with a single spiking microlaser

Double degree program:
Nanotechnologies for ICTs and Quantum Devices

Academic year: 2024/2025
Graduation session: July 2025

Supervisor:
Sylvain BARBAY

Candidate:
Federico ARNAUDO

Acknowledgements

I would like to express my deep gratitude to my supervisor Dr. Sylvain Barbay, for all the explainations and the the help he has given me during these months.

Alongside with him I want to also thank all my colleagues in the ToniQ group for helping me integrate as part of the team and for all the fruitful discussions.

I want also to thank all my colleagues, both from Torino and from the NANOQUAD program in Paris, for all the moments and the help they gave me in every moment of these two years in Italy and in France.

Finally I want to thank my family, my girlfriend Giorgia and my friends from home, for always being there to support and advise me both in the tough times and joyful moments.

Research context

The Master thesis has been carried out at Centre de Nanosciences et de Nanotechnologies, C2N, a joint research unit between CNRS (Centre National de la Recherche Scientifique) and Université Paris-Saclay, devoted to Photonics, Fundamental Physics, Material Science and Quantum technologies located in Palaiseau, Île-de-France, France.

More precisely, I have been working for four months in the ToniQ group, belonging to the Department of Photonics, that performs research in the fields of nonlinear and quantum optics, optomechanics, neuromimetic photonics, photonic crystals and photonic arrays of micropillars, waveguides and nanocavities.

The internship project has been carried out in the framework of neuromimetic photonics addressed by spiking micropillars. The broad research context of the group is to investigate the physics of these building blocks and how they could be used as nodes in a network to efficiently process information.

Table of Contents

List of Figures

Chapter 1

Introduction

Digital computers process information using discrete values, typically binary digits (0 and 1). In contrast, analog computers manipulate continuous variables, physically embodying the quantities being computed. While analog computing dominated computational technologies for centuries, the emergence of solid-state transistors and scalable digital architectures led to its widespread replacement. However, contemporary challenges such as machine learning, high-speed data processing and edge computing are driving a renewed interest in analog computing to overcome some of the limitations of purely digital systems [1]. In particular, the rise of artificial intelligence (AI) and large language models (LLMs) of the last few years underscores the increasing computational demands and energy costs associated with the training of advanced neural models.

1.1 Introduction to analog neuromorphic computing

The overwhelming majority of computers employ the Von Neumann architecture, in which data are stored in a dedicated memory (RAM) and transferred to a central processing unit (CPU) for processing, before being written back to memory. This architecture is inherently sequential and relies on continuous data transfers over a shared bus between the CPU and memory. This data movement introduces a bottleneck, as the processing speed of the CPU often exceeds the rate at which data can be transferred across the bus. Moreover, in recent years, the advancement of Moore's Law - which states that "the number of transistors incorporated in a chip will approximately double every two years" [2] - has slowed down due to power consumption constraints, further widening the gap between the computational efficiency of brains and conventional CPUs by orders of magnitude. Moreover, despite being outperformed by digital systems in terms of computational power, the human brain still presents many advantages in specific scenarios, such as facial recognition in low-light environments and speech understanding in presence of

noise. Finally, biological brains provide an evident robustness, since they continue to function even after a loss of some cells and do not require explicit programming or updates.

One promising approach to address this challenge is constituted by neuromorphic computing, which goes beyond software-based emulation of brain function (as in artificial neural networks), but also aims to develop new computing architectures inspired by the operational principles of biological systems, with the potential to replace or complement the conventional Von Neumann architecture.
The crucial task in the development of neuromorphic architectures is the design of artificial neurons, whose specific implementation strongly depends on the final application and on the structure of the system. In analog neurons weights are usually stored using resistors, charge-coupled devices, capacitors or floating gate EEPROMs.
However, together with the evolution of photonic solutions for optical communication and processing, optical implementations of neural networks have been a subject of investigation for a long time, mainly because of the high connectivity of optical system and the intrinsic parallelism offered through various degrees of freedom such as wavelength, polarization and mode. Specifically, different optical nonlinearities could be used to implement the characteristic neuronal functions. Although the first implementations were large and non-scalable, integrated photonics recently emerged as a key enabling technology, overcoming many limitations of early solutions. All these characteristics suggest that optical neural networks may surpass electronic solutions in computational speed and energy efficiency.

1.1.1 Reservoir computing

Most Artificial Neural Networks (ANNs) are structured into three key components. The input layer receives and feeds data into the network, which then connects to one or more hidden layers via a set of tunable weights. Finally, the last hidden layer is connected to the output where classification takes place. ANNs are predominantly divided in two categories: feed forward networks, where all the connections are directional from input to output, and recurrent neural networks (RNNs), where there are feedbacks within single nodes and interconnections within the nodes of each hidden layer are allowed. These architectures have demonstrated extensive capabilities in processing static or, in the case of RNNs, complex temporal tasks. However, practical applications require to overcome challenges such as high computational costs and slow convergence of training algorithms.

To surpass these obstacles, the concept of Reservoir Computing emerged, at first with the principles of Echo State Networks (ESNs) [3] and Liquid State

Machines (LSMs) [4], and later with the presence of a nonlinear physical device [5]. These configuration make use of a random, recurrent network architecture where the internal weights remain untrained, significantly reducing the training process by only adjusting the output weights. The generic structure of a standard reservoir computer, illustrated in Figure 1.1(a), includes three neural layers: an input layer, a reservoir containing discrete neurons and a readout layer. In this configuration the reservoir here acts as a dynamic filter, exploiting its dynamical properties to transform input signals through a high-dimensional temporal mapping, thereby facilitating the processing of temporal information. In the physical reservoir framework, the internal weights must be accurately scaled to make the physical device(s) that constitute the reservoir hover near a point of dynamical instability. The attractiveness of the physical implementations of RC is further magnified by its vast adaptability to a vast range of nonlinear system, making it a powerful and versatile tool for research and physical implementation of neuromorphic computing schemes. Some examples are reported in section 1.2.

1.1.2 Spiking Neural Networks

The physical system presented in this work is based on a reservoir computing architecture, which is then adapted for Spiking Neural Networks (SNNs).
SNNs constitute the closest class, among all the ANNs, to biological systems and have gained attention in several biosensing areas, including image and sound processing or fast motion control [6].

In SNNs, whose generic structure is represented in Figure 1.1(b), information is encoded both in the digital and analog characteristics of spikes. Furthermore, by leveraging the inherent synchronization properties provided by these systems - and by explicitly taking into account the precise timing of spike events - this architecture is expected to offer greater computational power compared to classical ANN models based on sigmoidal neurons.

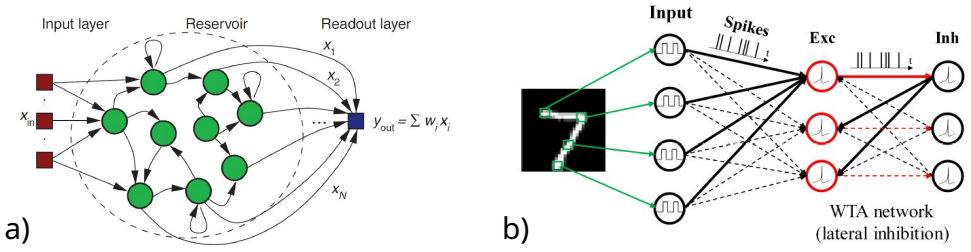

Figure 1.1: General schemes of RC (a) and SNN (b). Figures adapted from [1] and [7]

The artificial neurons that are employed in spiking neural networks aim to

mimic their biological counterpart, and therefore need to employ some of their characteristic features:

Excitability It refers to the ability to respond to external stimuli by generating an all-or-none response, known in the biological field as *action potential*. The response must be well calibrated (or even independent) with respect to the applied stimulus. The nonlinear function describing the relationship between the synaptic input received by a neuron and the generation of an action potential is know as *activation function*, which can be often modeled as a sigmoid or a rectified linear unit.

Refractory period It consists in a transient time after the excitation of an action potential during which the neuron is unable (absolute refractory period) or less able (relative refractory period) to generate another action potential. In the second case, a stimulus can still be created by stronger-than-usual stimuli.

Summability It refers to the ability of the neuron to integrate multiple incoming synaptic inputs to determine whether to fire a spike. The integration can occur through both temporal and spatial summation. The first takes place when consecutive inputs from the same synapse arrive within a short time window, causing the postsynaptic potential to accumulate. The latter, instead, involves the simultaneous arrival of stimuli from multiple synapses distributed across the dendritic tree, whose combined effect can bring the activation function over the firing threshold.

Plasticity It refers to the capacity of a neuron's structure and function to change over time in response to its activity, experience or environmental influences. This adaptability, which can involve modifications of synaptic strength or alterations in the activation threshold, may manifest as a short-term plasticity, lasting milliseconds to minutes, or as long-term plasticity, resulting in persistent changes that can endure for even years.

1.2 Overview of neuromimetic photonic systems in literature

Numerous semiconductor-based photonic systems have been investigated to emulate the behavior of biological neurons. In 2011, Appeltant et al. [8] demonstrated both theoretically and experimentally that information processing can be achieved using a single dynamical node with an optical feedback loop in a RC physical architecture. Chakraborty et al. proposed a simulated neural network composed of

all-photonic spiking neurons based on phase-change materials [9]. In 2017, Shen et al. developed an integrated silicon photonic feed-forward neural network, both theoretically and experimentally, employing a cascade of Mach-Zehnder interferometers to harness the properties of coherent light [10]. Extensive research activity is also being performed in the framework of photonic synapses [11], which would be responsible for the collection and transmission of signals through the neurons in a real neuromimetic architecture.

In 2011, the group of S.Barbay proposed a physical implementation of a spiking neuron based on a semiconductor laser with a saturable absorber [12]. This approach is particularly appealing as it removes the need for optical feedback loops, thereby simplifying the experimental setup. Additionally, it avoids the use of slow thermal mechanisms typically involved in phase-change materials.

Building on the PhD thesis by A. Masominia in the group [1, 13], the goal of this internship is to push further the numerical and experimental investigation of the computational potential of such system, with the ambitious goal of developing a fully optical ultra-fast classification scheme.

In Chapter 2 I will describe the physical system that I am going to use as an artificial neuron and I will give some numerical results regarding the general neuromimetic properties of a single micropillar.
In Chapter 3, instead, I will present numerical simulations addressing the whole process of image classification, along with preliminary experimental results characterizing the properties of the system.

Chapter 2

Physical system

2.1 Neuromimetic micropillar laser

As previously discussed, in recent years, neuromimetic photonic systems have gained substantial attention owing to their promise for implementing analog, spike-driven, and non-Von Neumann computational paradigms [14]. These systems aim to replicate a central feature of biological neurons, that is their excitability, which allows for the encoding of information into sharp and well-defined spikes that facilitate both signal transmission and computational functions.

Lasers with saturable absorbers[1] have been extensively studied as excitable systems, both from a theoretical [15] and an experimental [16] point of view.

While the majority of the proposed devices required coherent laser injection into a cavity, a major breakthrough came along with the development of the first excitable semiconductor laser incorporating an integrated saturable absorber [12, 17]: this kind of system needs only incoherent pumping, either electrical or optical, and its functionality is based on a carrier based slow-fast non-linearity instead of slow thermal responses. Moreover, the micropillar lasers can be easily integrated into arrays and other spatial schemes exploiting their evanescent coupling [18].

2.1.1 Excitable microlaser

The excitable system considered during the internship, represented in Figure 2.1, is a specifically designed, optically pumped VCSEL featuring an integrated intracavity saturable absorber. The monolithic cavity incorporates aperiodic multilayer AlGaAs/AlAs mirrors engineered to obtain efficient optical pumping

[1] A saturable absorber may be generically defined as a material whose absorption decreases as the intensity of the incoming light increases. At low intensities, the material strongly absorbs light, but as the intensity rises, the absorption "saturates," and the material becomes increasingly transparent.

across a wide wavelength spectrum [19]. The active zone of the device contains two InGaAs/GaAs quantum wells for the gain medium and one InGaAs/AlGaAs quantum well acts act as a saturable absorber. All the three quantum wells have to be placed at the anti-nodes of the cavity resonant fields, targeted at 980 nm. Moreover, in order for it to remain unpumped, the placement of the SA quantum well must be engineered to correspond also with a node of the pump field, which is characterized by a wavelength of 808 nm.

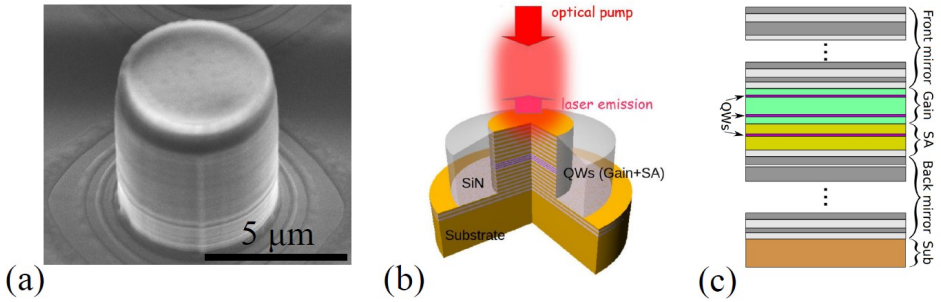

Figure 2.1: Real and schematic images of the micropillar. (a) Image of a micropillar laser. (b) Structure of the excitable VCSEL. (c) Simplified scheme of the device architecture.

The physical working principle of the device is as follows: when an optical pump is applied to the system, electrons in the valence band of the gain quantum wells are excited, leading to a large increase in the laser gain (defined as $G(\nu) = \sigma(\nu)\Delta N$, with $\sigma(\nu)$ the stimulated emission cross-section and $\Delta N = N_c - N_v$ the difference between the carriers' concentrations in the excited and fundamental states). Meanwhile, the saturable absorber (SA) remains unpumped due to its spatial location at a node of the pump field. The resulting spontaneous emission processes that emerge from the pumped gain region are responsible for absorption processes inside the SA, which effectively block most of the resulting light intensity. However, once the intensity reaches the saturation threshold of the absorber, its absorption decreases and the SA becomes transparent to light. This allows a rapid onset of stimulated emission, which quickly depletes the carrier population in the gain region, resulting in a macroscopic light spike. In a comparable timescale, the excited electrons in the conduction band of the SA relax back to the ground state. Under suitable conditions—further detailed in Section 2.1.2—this process repeats periodically over time, typically at frequencies in the gigahertz range [20].

2.1.2 Yamada model

The micropillar excitable dynamics is described by a modified version of the well-estabilished Yamada model [21], which is described by the following set of non-linear, coupled and adimensional differential equations, that can be derived from the rate equations for the system, taking into account absorption and stimulated emission processes, non-radiative recombinations and a spontaneous emission term for the intracavity intensity (neglected in the remaining two equations):

$$\begin{cases} \dot{I} = I(G - Q - 1) + \beta_{sp}(G + \eta)^2 \\ \dot{G} = \gamma_G(\mu_1(t) - G(1 + I)) \\ \dot{Q} = \gamma_Q(\mu_2 - Q(1 + sI)) \end{cases} \tag{2.1}$$

These equations characterize the intra-cavity intensity I, while G and Q are, respectively, proportional to the excess carrier densities in the gain and saturable absorber regions. The factor $R(t) = G(t) - Q(t) - 1$ is defined as the rescaled net gain. Different parameters influence the system dynamics: μ_1 and μ_2 represent respectively the pump intensity and the linear absorption, η adjusts for the transparency offset in the gain region, γ_G and γ_Q are the non-radiative recombination rates, s is a saturation parameter that takes into account the finite number of states in the conduction band of the SA and β_{sp} is the spontaneous emission factor, which accounts for the fraction of spontaneously emitted photons that effectively couple into the cavity's resonant mode. Across the simulations, some of the previously presented parameters remained fixed; their numerical values are reported here: both γ_G and γ_Q have been set to a value of 0.005, $\mu_2 = 2.0$, $s = 10$, $\eta = 1.4$ and $\beta_{sp} = 10^{-5}$. All the factors of Eq. 2.1.2 are rescaled with respect to the photon lifetime in the cavity; therefore, a time unit in the system corresponds approximately to 1.5 ps.
As one may notice from Eq. 2.1.2, the effective timescales governing the dynamics of G and Q are on the order of $\tau_{G,Q} = 1/\gamma_{G,Q} = 200$, whereas the evolution of I occurs on a much faster timescale, approximately $\tau_I \sim 1$. This separation of timescales gives rise to a slow-fast nonlinear system, where G and Q act as slowly varying variables, while I responds rapidly to changes in the system.
In order to maintain physical coherence with the experiment during this study, perturbations will be applied on the pump μ_1, since it will be the only parameter that may be accessed through the experimental setup.

Figure 2.2 represents the bifurcation diagram evaluated from the system characterizing the Yamada model.

The graph represents the maximum value reached by the intra-cavity intensity I as a function of the base pump μ_1. For low values of μ_1 the system is set in a

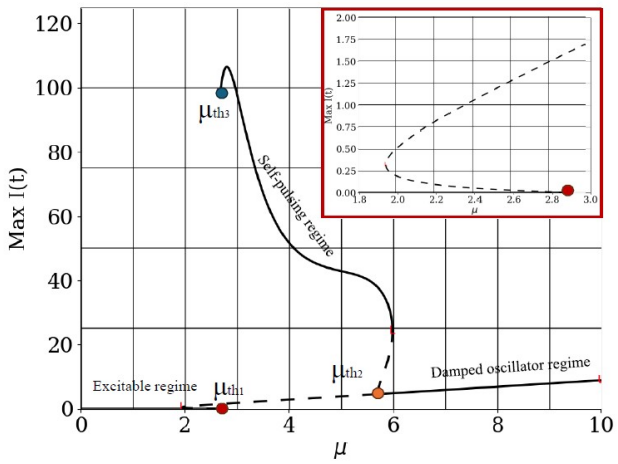

Figure 2.2: Bifurcation diagram of the solution of the Yamada model; solid lines represent the stable solutions of the system, while dashed lines the unstable ones. Image adapted from [1].

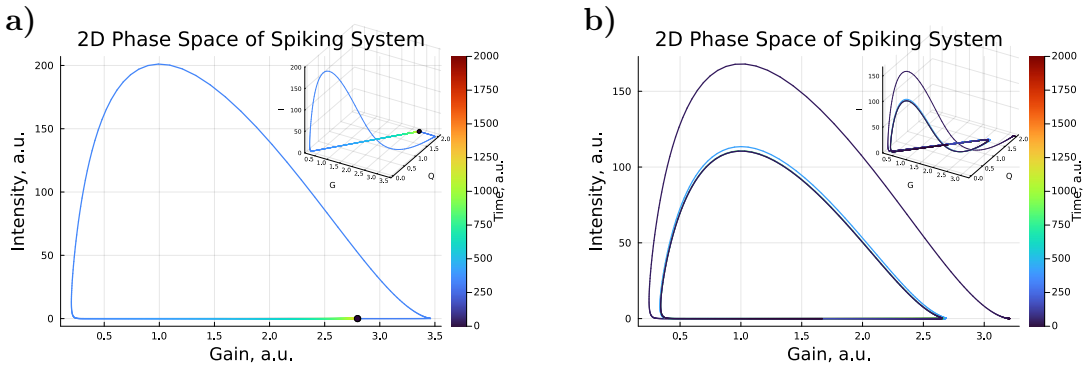

Figure 2.3: Projection on the G–I plane of the phase representation of: **(a)** the emission of a spike after a perturbation. After spiking, the system returns to the stable state; **(b)** self-pulsing regime, where spikes are emitted continuously. Insets show the full phase space trajectories considering all three variables.

non-lasing state: the stable solution is $I = 0$. By following the maximum amplitude in the graph, an abrupt jump is then clearly visible after the excitable threshold μ_{th1}: when the threshold is surpassed, the system immediately starts to emit light pulses of high intensity without any variation in the pump intensity. If μ_1 is further increased, the intensity of the self-spiking pulses decreases until a second jump is found. For high values of pump, the microlaser starts emitting a continuous wave signal (in the image denoted as "damped oscillator regime"), with an intensity that is notably lower than the one coming from the self-spiking regime. However, the

most interesting regime for tasks such as image classification is the excitable regime (the one right under μ_{th1}), in which a spike is emitted when a perturbation of the pump is enough to make it overcome the spiking threshold and, therefore, to emit a spike.

Additionally, Figure 2.3 shows the phase-space representation of the intensity I as a function of the gain G both in the case of a single excitable spike (a) and of a self-spiking regime (b). Clearly in the second case the system converges toward a limit cycle corresponding to the emission of a spike train with a definite period. The black dot in Figure 2.3(a) represents the stable solution of the system, to which the system converges after the emission of a spike.

2.2 Numerical results

In the preliminary computational studies that will be presented in this section, the system of three coupled ordinary differential equations (ODEs) has been integrated by using is the `solve_ivp` function from the Python library *SciPy* [22]. This function, which interfaces to different ODEs solvers, is able to manage both stiff and non-stiff problems with adaptive step size control; this adaptability is essential for capturing the complex slow-fast dynamics of the system. The explicit Runge-Kutta method of order 5(4) 'RK45' has been specifically selected for the simulation of the system.

2.2.1 Micropillar's neuromimetic properties

The objective of this section is to numerically characterize the properties of the excitable system introduced in Chapter 1, that enable the device to emulate the behavior of a biological neuron. In order to maintain a coherence with the experimental work that will be described in section 3.3, in order to demonstrate all the following properties the pump perturbation applied to the system has a boxcar shape selected according to the following formula:

$$\mu_1(t) = \mu_{1,0} + p\Pi_{\tau_p}(t - t_0) \tag{2.2}$$

in which $\mu_{1,0}$ is the base pump of the system and represents the pump intensity needed to bring the system in the excitable regime, while the second part of the equation represents a boxcar function with amplitude p and time width τ_p; t_0 is the initial time at which the perturbation is applied.

Excitability The key property that a neuromimetic spiking microlaser must exhibit is excitability. As shown in Figure 2.4(a), the first, weaker perturbation of the pump fails to trigger a spike, whereas the second, stronger perturbation successfully elicits a macroscopic response. A more detailed analysis reveals that the critical factor governing excitability is the net gain. In fact, whenever the net gain $R(t)$ exceeds the threshold defined in Section 2.1.2, a spike is almost instantaneously generated.

Figure 2.4(b) further emphasize the excitable threshold of the system by plotting the peak intensity reached during the simulation as a function of the perturbation amplitude. As clearly represented, the over-threshold spike amplitude is well calibrated and linearly follows the amplitude of the perturbation. Figure 2.4(c) , instead, demonstrates the self-spiking regime, in which the system emits periodic spikes spontaneously once the baseline pump exceeds the threshold, even in the absence of external perturbations. The sharp transition between the two different regimes is a crucial feature for information processing. Although it is theoretically possible to generate light intensity spikes by using short gain switching - i.e. a high-intensity pulse in the laser pump - in a conventional laser, the linear response of such laser means that most of the information encoded in weaker smaller stimuli would be lost.

Figure 2.4(d) shows the excitability threshold, defined as the minimum value of p (i.e. the maximum of $1/p$), required to trigger a spike as a function of the perturbation width τ_p in Eq. 2.2. It can be clearly observed that the energy of the pump perturbation, $E = \delta\mu_1\tau_p = p_i\tau_p$, necessary to elicit a spike, remains constant for $\tau_p \gg \frac{1}{\gamma_{Q,G}}$.

However, for large τ_p, the curve saturates because the system compensates the carrier injection from the pump with increased losses due to non-radiative recombination processes [1, 23].

Refractory period As shown in Figure 2.5(a), on the left side of the figure, the proposed neuromimetic system clearly exhibits an intrinsic absolute refractory period: although perturbations carry sufficient energy to trigger a spike, only one spike effectively does so. On the right side, instead, the presence of a relative refractory period is highlighted, meaning that both perturbations elicit responses, but the second is notably weaker. This reduced response results system's incomplete return to equilibrium, due to the slow recovery dynamics of the gain (G) and saturable absorber (Q). The refractory period of the spiking micropillar naturally correlates the temporal sequence of incoming stimuli, effectively acting as form a short-term memory. This enables the use of a single spiking micropillar as a complete physical reservoir, eliminating the need for a feedback loop that would otherwise be required.

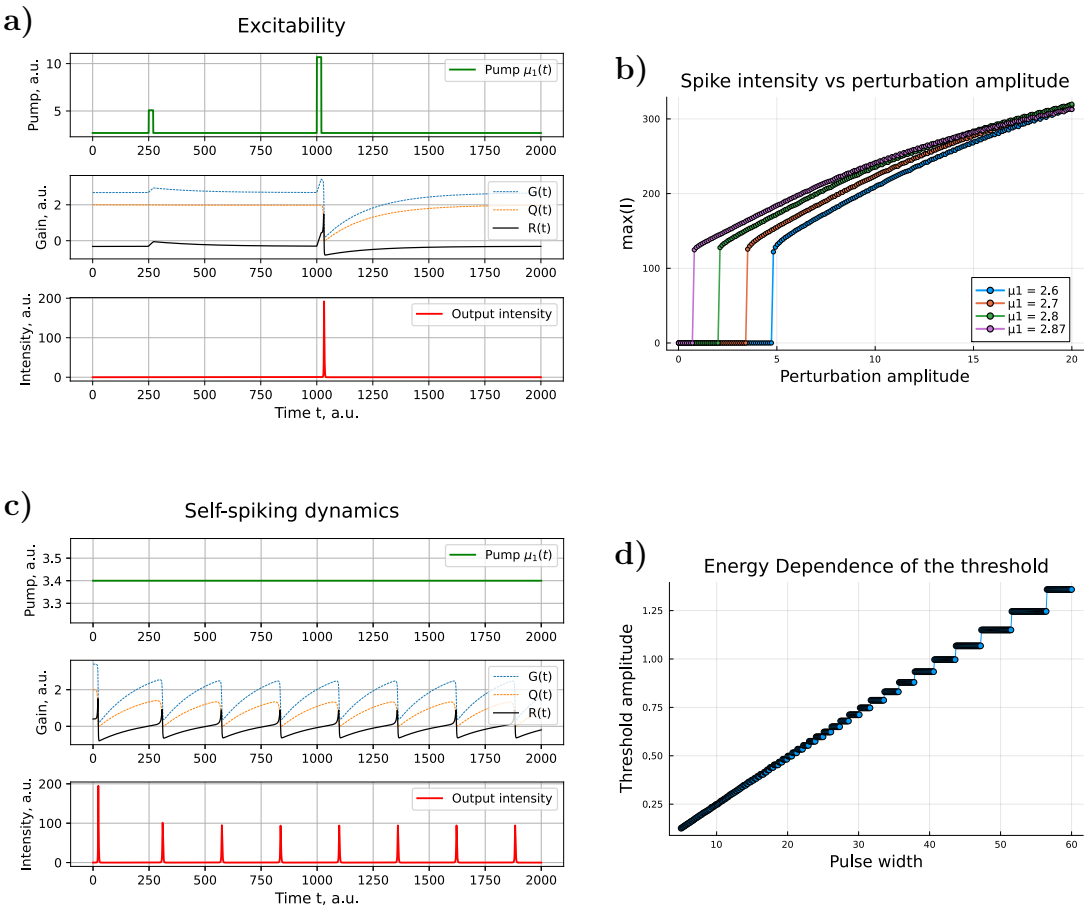

Figure 2.4: Numerical demonstration of the excitability of the system: (a) Only perturbations that make the net gain exceed a threshold trigger a spike; (b) Maximum value of the emitted intensity as a function of the perturbation amplitude. The minimum perturbation needed to excite a spike depends on the base pump; over the excitable threshold the spiking intensity is clearly reproducible. (c) A base pump $\mu_{1,0} \geq 3.0$ causes self-spiking. (d) Inverse of the minimum perturbation amplitude required to excite a spike as a function of the pulse width; for high values of the pulse width, the graph shows a step-like structure caused by numerical limitations, as it was obtained via brute-force physical simulations across numerous p_i and τ_p pairs, rather than through a continuation algorithm.

Temporal summation As illustrated in Figure 2.5(b), contributions from multiple pulses in the laser pump can sum temporally - provided they occur close enough in time - to collectively trigger a spike. This behavior is essential for the biomimetic character of the device.

The combination of these properties makes the system particularly well-suited for implementation in hardware neural networks and also enables a direct recognition of simple patterns, as highlighted by Masominia et al. in [13]. For instance, Figure 2.5(c), demonstrates that, with proper tuning, the system can distinguish between closely related input sequences such as 1011 and 1101. This capability can be used to design a simple photonic spike-based classification scheme [1].

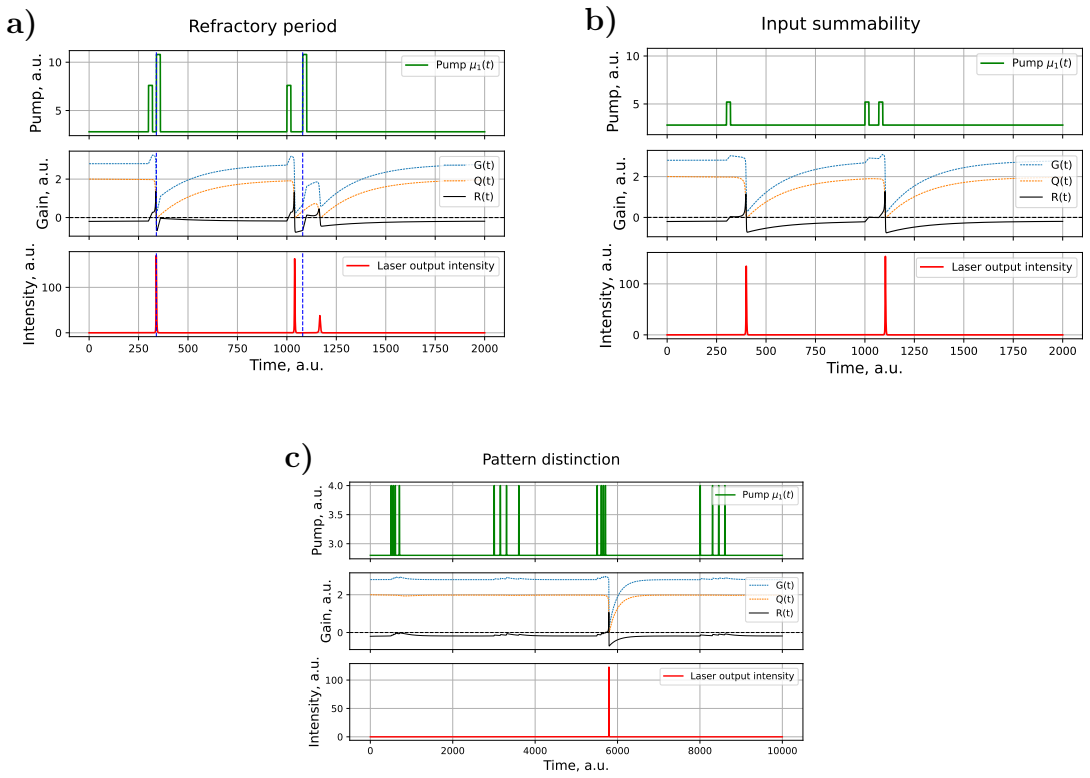

Figure 2.5: Properties of the excitable system: (a) Absolute (left) and relative (right) refractory period of the system; the blue dashed lines indicate the moment in which the second perturbation is applied. (b) (c) A correct tuning of the parameters allows the distinction of similar features such as "1011" and "1101"

.

Spike latency The system exhibits a characteristic delay - here referred to as *spike latency* - between the application of a perturbation on the pump and the spike coming from the system's response.

This latency arises from the slow-fast dynamics governing the excitable behavior of the system: when a perturbation is applied, the system does not respond instantaneously; instead, the response is delayed by an amount of time that depends

sensitively on the energy of the perturbation relative to the excitable threshold. More in detail, as represented in Figure 2.6(a), if the net gain barely exceeds the threshold, the system "hesitates" near the resting state before transitioning to an excitable response, resulting in a long latency. In contrast, perturbations that make $R(t)$ exceed the threshold by a larger margin lead to shorter delays. The latency dramatically increases as the input energy approaches the threshold from above, as illustrated in Figure 2.6(b).

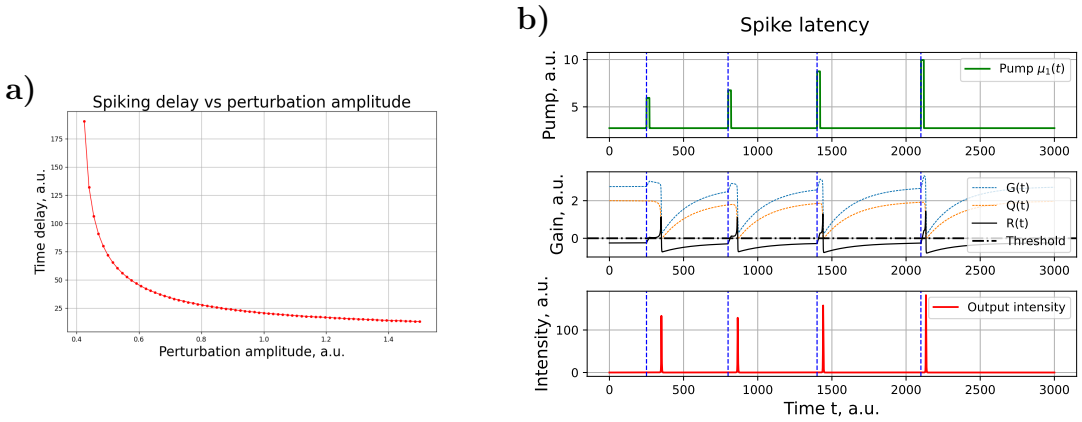

Figure 2.6: Spike latency of the system: (a) Spike latency as a function of the input perturbation amplitude. (b) Examples of the response of the system to different input perturbations.

This property paves the way for the enrichment of the information carried by individual spikes. As proposed in [13], indeed, it would be possible to make use of the difference between the time delays of the spikes excited by different perturbation patterns, thus employing a *temporal coding* instead of the more conventional event-based coding, to increase the computational efficiency for tasks such as image classification.

Chapter 3

Image classification with spiking microlasers

The broad objective of the work is to pave the path toward a fully hardware-based computing system. Although the initial studies that will be described in this chapter do not eliminate the need for a conventional computer, they still represent an important step to comprehend the mechanisms of encoding and decoding of the data. Exploiting the information presented in chapter 1, in this chapter the spiking microlaser will be exploited as a nonlinear node within a reservoir computing framework under the hypothesis that the sparsity and event-based nature of the spikes may facilitate the mapping of the data into an higher dimension: each spike is elicited under specific conditions and conveys substantial information about the perturbation that induced it.
While conventional reservoir computing architectures typically rely on a complex network of recurrently connected nonlinear nodes to retain memory of past inputs, in this case of study a single spiking microlaser can effectively replace the entire recurrent network. This is made possible by the intrinsic refractory period of the microlaser, which provides a form of short-term memory by temporally correlating the system's response to recent stimuli. As a result, the need for explicit recurrent connections is relaxed, allowing the memory of the system to emerge from its own dynamical properties, thus enabling the system to retain memory through its intrinsic dynamical behavior rather than through explicit recurrent connectivity.

Image classification is a standard benchmark for evaluating the performance of machine learning algorithms. In particular, digit recognition represents an interesting case: while it does not require highly complex models, it still presents enough challenge to serve as an intermediate step toward real-world applications in computer vision and pattern recognition.

3.1 Methodology

As discussed in section 2.1.2, pump level (μ_1) of the micropillar gain is able to set the system in the so-called excitable regime so that a spike is excited only when a certain perturbation is applied to the system. This condition is essential for the image classification tasks. In this context, the system must undergo both a training phase (sec. 3.1.1), where it has to "learn" the association between input patterns and output labels, and an inference phase (sec. 3.1.2), where it applies the learned behavior to classify new inputs. In the first phase, the images have to be encoded in a suitable spatial or temporal scheme in order to be processed by the neural network. In the context of reservoir computing, the output of the reservoir is decoded in a vector and then used to evaluate the optimal readout matrix. During the inference phase, instead, the decoded ouptut is multiplied for the evaluated matrix. The result is a vector containing all the "scores" assigned to each label.

Dataset

The reduced MNIST digit dataset provided by scikit-learn [24] consists of 1,797 grayscale images of handwritten digits, each in an 8×8 pixel format. Every image is flattened into a 64-dimensional vector, where each element represents a pixel intensity ranging from 0 (black) to 16 (white), across 17 discrete grayscale levels. The dataset is designed to be balanced, containing approximately 179 or 180 samples for each of the ten digits (0–9). In the studies described, if not specified, 70% (125 images per digit) of the dataset has been employed for the training and 30% (54 per digit) for the inference analysis and accuracy evaluation of the system. In order to avoid the possibility of casually selecting a particularly well-suited training subset, measurements are repeated 10 times, each with a new randomly selected set of digits.

Solver

The numerical computations were performed using the *Julia* programming language, known for its high performance in scientific computing. In particular, the integration of the differential equations was carried out using the *Tsit5* solver, an explicit Runge-Kutta method of order 5(4), which is well-suited for non-stiff problems.

3.1.1 Training

In the framework of reservoir computing, which this work is based upon, training the system means to compute the optimized readout matrix. The algorithm used for the training of the model is reported in Figure 3.1

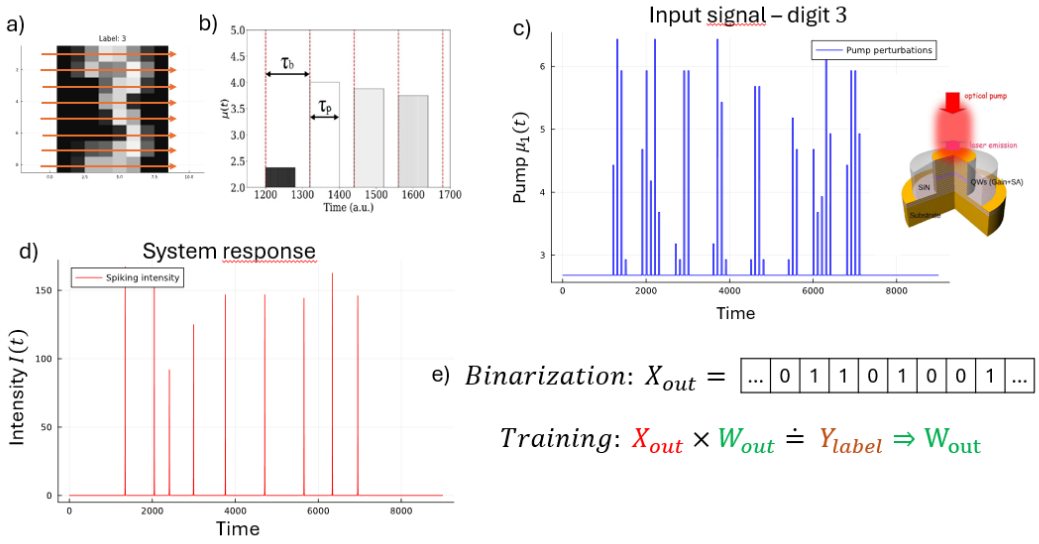

Figure 3.1: Training algorithm used in the simulations. (a) Example of one image from the dataset in 8x8 format. (b) Schematic of the bit time (τ_b) and perturbation time (τ_p). (c) Encoding process: the digit vector is converted into a pump perturbation scheme. (d) Optical response from the system. (e) The response is decoded into a binary vector and the linear system is solved to find W_{out}

Encoding

The first step for the training of the classification model is the encoding of the images into a specific perturbation scheme: in this case the perturbations are applied on the pump of the excitable laser. Each dataset entry (64 values, one for each pixel of an image) are stored as a vector of 64 values.
In order to generalize the problem, a random vector with the same dimension is created for each digit that is inputed to the system and a dot product is performed. In contrast to conventional machine learning algorithms, during the simulations we do not enlarge the data volume through masking procedures. Instead, we aim to fully exploit the information already contained in the dataset.

The resulting vector containing the perturbation amplitudes p_i is then normalized and converted using a coefficient c into the pump laser $\mu_1(t)$ according to the following formula:

$$\mu_1(t) = \mu_0 + \sum_i c p_i \Pi_{\tau_p}(t - i\tau_b) \tag{3.1}$$

where $\Pi_{\tau_p}(t)$ is a boxcar function of duration τ_p, μ_0 is the base (DC) pump, which is used to keep the laser in the excitable regime and is kept constant during the

simulation. τ_p and τ_b are the respectively the perturbation and bit times, which are clarified in Figure 3.1 (c): each pixel perturbs the system for a time τ_p, and occupies an effective time τ_b. The factors in this formula are the actual degrees of freedom that can be modified in the system.

In the simulations, the N digits are fed sequentially into the spiking micropillar by appending a (1, 64) vector to the input matrix for each image provided. This results in an input matrix referred as $X_{in} \in \mathbb{R}^{N\times 64}$.

Decoding

Once the input is provided to the system, the response consists of a train of spikes whose physical characteristics—such as amplitude and temporal spacing—depend on the input. Due to the intrinsic refractory period of the photonic neuron, not all pump perturbations are converted into spikes. This results in a sparse spiking output, which enhances the information content of each individual spike.

To compute the optimized readout matrix, the resulting analog signal must be converted into digital form. This is achieved by dividing the entire response timeline into a fixed number K of time bins and assigning a binary value to each bin based on whether the light intensity within that interval exceeds a given threshold. In this work, the threshold is set to 20% of the image's maximum intensity, although it can be further optimized as a hyperparameter. A representation of this procedure is represented in Figure 3.2. As for the images, all the resulting output vectors are concatenated in a matrix $X_{out} \in \{0,1\}^{N\times K}$.

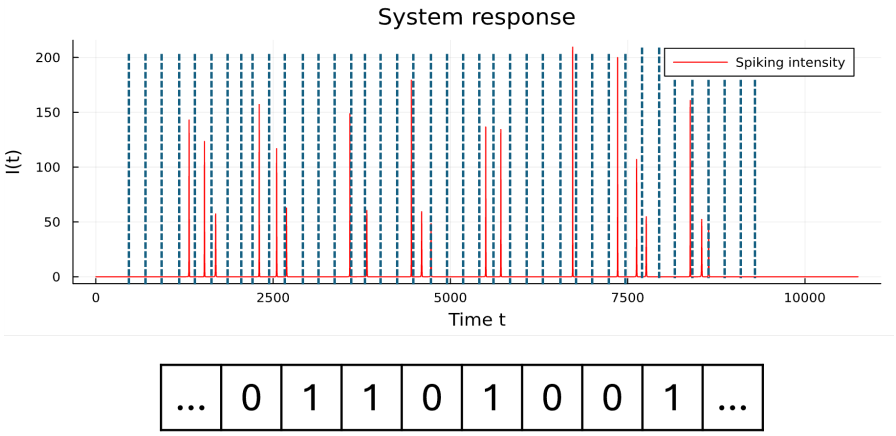

Figure 3.2: Decoding of the output: by considering a number K of time bins, each bin will take the value 1 in the presence of a spike and a 0 in its absence

Training of the readout matrix

As already described, in the framework of reservoir computing, on which this work is based upon, training the model refers just to optimizing the readout matrix $W \in \mathbb{R}^{K\times 10}$. In machine learning literature this can be assessed in a variety of possibilities [25–27]. However, in order to pursue the broad goal of the project of minimizing - and ideally eliminating the system's reliance on an external computer - we adopt the simplest form of supervised training. This consists in solving the linear system $XW = Y$, where $Y \in \{0,1\}^{N\times 10}$ contains the one-hot encoded labels of the training dataset. Since the system is overdetermined ($N > K$), this has to be done by looking at the solution that minimizes the squared error $\|XW - Y\|_F^2$, where $\|\cdot\|$ is the Frobenius norm. The resulting least-squares solution, which effectively computes the best linear mapping, is then

$$W = X^+ Y \tag{3.2}$$

in which X^+ is the Moore-Penrose pseudoinverse of X. This approach provides a closed-form solution to the problem, without requiring iterative optimization methods such as gradient descent.

3.1.2 Inference

Inference is the stage in which a previously trained machine learning model is deployed to generate predictions on data it has not encountered before. Once the model parameters - in this case the readout matrix W - have been optimized during the training phase, the model applies the learned input-output relationship to new examples.

During this phase, the system takes an input sample $x_{in} \in \mathbb{R}^{1\times 64}$, that is encoded as during the training phase described in section 3.1.1. The decoded output $x_{out} \in \{0,1\}^{1\times K}$ is then multiplied by the optimized matrix $W \in \mathbb{R}^{K\times 10}$ found during the training: the result is the vector $Y_{pred} \in \mathbb{R}^{1\times 10}$ that contains a "score" for each label, that is associated to the pseudo-probability of each label to be associated to the image x_{in}. The predicted label will be the one associated to the highest score.

By testing the model with about 30% of the dataset one can finally compute the accuracy of the system as

$$A = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} = \frac{1}{n}\sum_{i=1}^{n} I(\hat{y}_i = y_i) \tag{3.3}$$

where I is the indicator function, the value of which is 1 if the prediction $\hat{y}_i$ is equal to the true label y_i and 0 otherwise.

3.2 Numerical results

In order to test the capabilities of the described system, at least from a numerical point of view, different simulations have been carried out by considering the dataset of handwritten digits described in section 3.1.

For the initial analysis, a dataset composed by 125 samples per digit is used. Both training and inference are repeated 10 times with the digits shuffled each time to avoid bias from randomly select a favorable dataset.
Several parameters are kept fixed across all the simulations presented in this work: $\mu_2 = 2.0$, $\gamma_G = \gamma_Q = 0.005$, $S = 10.0$, $\eta = 1.4$, $\beta_{sp} = 10^{-5}$, $\tau_p = 10.00$. The remaining parameters of Eq. 3.1, instead, are considered as hyperparameters subject to further optimization. During the decoding phase, the data are binarized using a threshold set at 20% of the maximum output intensity produced during the training. Additionally, the output spike stream is divided into time bins equal in number to the pixels in the image, which is 64. Figure 3.3(a) shows an example of the pump perturbation scheme applied to the microlaser and the corresponding spiking output.

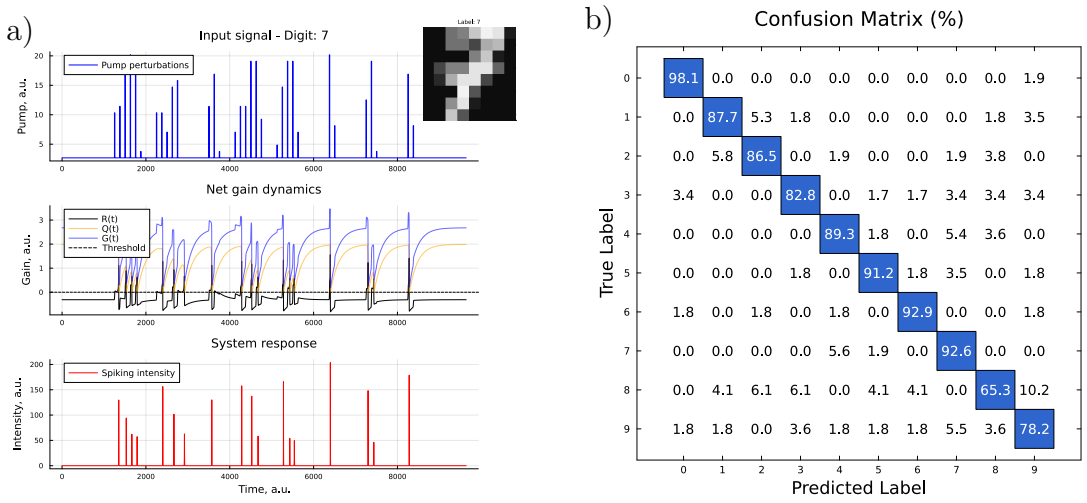

Figure 3.3: (a) Top: pump perturbation scheme given as input to the micropillar; center: net gain dynamics of the optical neuron, the net gain acts as a thresholding function; bottom: response of the excitable laser to the input; inset: example of one of the images in the MNIST dataset. Pixels are read in the conventional order: left to right, top to bottom; parameters: $\mu_{1,0} = 2.68$, $\tau_b = 125.0$, $c = 17.5$. (b) Confusion matrix, the diagonal represents the correctly assigned labels.

As expected, due to the system's refractory period, not every input perturbation triggers a spike in the output. Furthermore, the amplitude of some spikes is significantly reduced in comparison to the expected values because of the relative refractory period that the system presents. Finally, Figure 3.3(b represents the confusion matrix, in which the predicted labels from the model are compared to the true labels of the dataset, that has been obtained after testing the system with the remaining 30% of the dataset that was not used during the training. For this reason, we opted to exploit the ultrafast properties of the optical scheme not to compute new data from the initial ones, but to physically read each image with different orders, as specified in section 3.2.1

3.2.1 Data augmentation

A common practice in machine learning is to increase the information content of a dataset by applying data augmentation techniques. With the training algorithm considered for this experiment (sec. 3.1.1), one might simply perform a matrix multiplication between the input matrix $X_{in} \in \mathbb{R}^{N\times 64}$ and a weight matrix $W_{aug} \in \mathbb{R}^{64\times K}$; doing so the dimensions of the readout matrix W_{out} remain coherent with the considered algorithm and independent from the size of the input dataset. However, as already stated in this work, one of the aims of the study is to minimize the dependence from an external computer to finally create fully optical setup for image classification. This approach comes from the notion of receptive fields in ANNs [28],

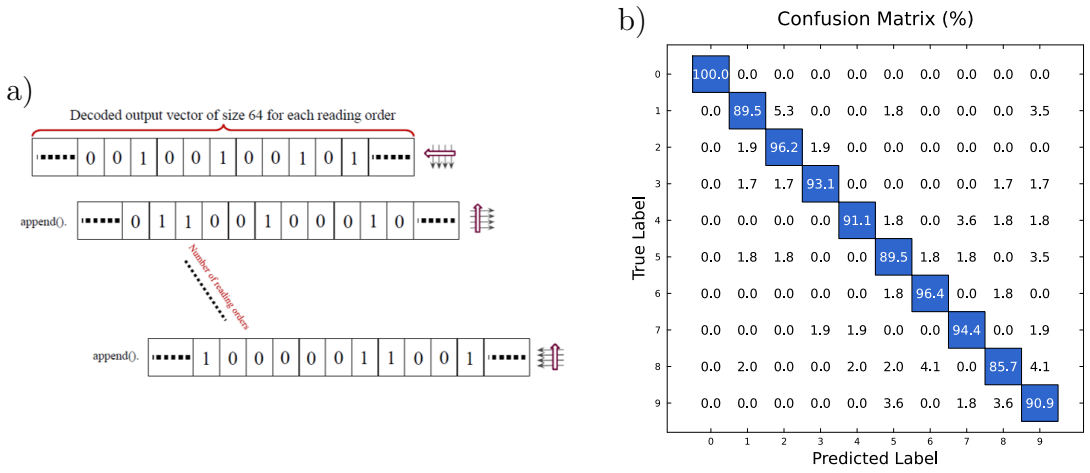

Figure 3.4: (a) Decoding process of the reading order method with a varying number of considered orders. (b) Resulting confusion matrix using three different reading orders; $\mu_{1,0} = 2.68$, $\tau_b = 125.0$, $c = 17.5$

that exploit the correlations between the data encoded in different variations or portions of each image to enrich the available information. In this configuration, each different reading of an image has to be horizontally appended to the input matrix, which therefore becomes $X_{in} \in \mathbb{R}^{N\times(64\cdot N_c)}$, where N_c represents the number of reading orders considered during the training. This process is represented in Figure 3.4. Obviously, the number of time bins has to be opportunely adjusted taking into consideration the amount of different reading orders considered. By considering all the eight possible horizontal and vertical reading orders, the simulation returns an accuracy $A = 94.14 \pm 0.96\%$. However, by performing different simulations, one can clearly notice how the accuracy saturates by considering more than three reading orders. By considering, as an example, only the "left to right, top to bottom" , "top to bottom, right to left" and "left to right, bottom to top" orders, the average accuracy found from the simulations is $94.07 \pm 1.09\%$. Therefore, all the following analysis will be performed using these three reading orders, which result in a matrix $X_{in} \in \mathbb{R}^{N\times 192}$

3.2.2 Dependencies and optimization

Several parameters in the proposed model influence the simulation results and can be further tuned.
The first - and arguably the most critical from a computational standpoint — is the number of time bins, which directly determines the size of the output matrix to be optimized. It is worth noting that, although the current model operates in a

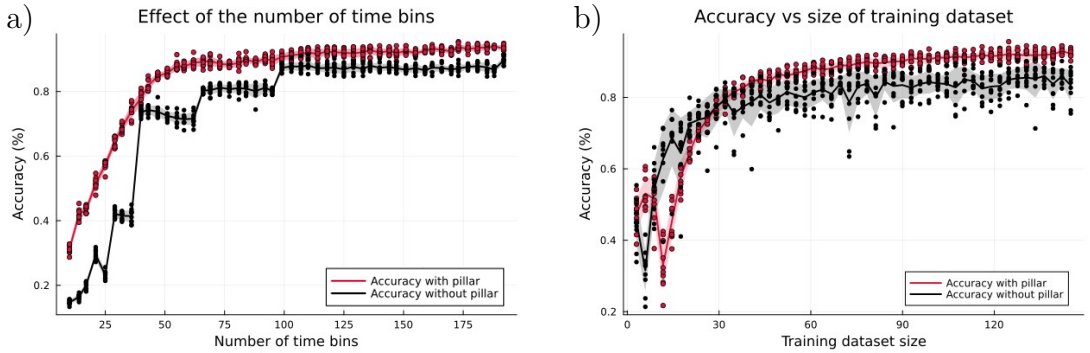

Figure 3.5: (a) Relation between the accuracy of the model and the number of time bins/virtual neurons considered. 125 samples for each digit were used for the training; parameters: $\mu_{1,0} = 2.68$, $c = 17.5$; three reading orders were considered for each digit. (b) Accuracy dependence with respect to the size of the training dataset, expressed as the number of samples given for each different digit; parameters: $\mu_{1,0} = 2.68$, $\tau_b = 125.0$, $c = 17.5$; three reading orders were considered for each image and the number of time bins was set to 192.

serial fashion by feeding all perturbations corresponding to each training image into the same node sequentially over time, an alternative approach would be to employ multiple nodes working in parallel; this would reduce the computational complexity at the expense of a more complex architecture. In a complete parallel configuration, the number of required nodes would match the number of time bins used in the serial case.

Figure 3.5(a) reports the dependence of the resulting accuracy with respect to the number of time bins considered during the simulation. To evaluate whether the neuromorphic system implemented in this work offers any advantage over a purely software-based solution using the same resources, the accuracy of the model can be compared to that obtained by directly computing W_{out} from the dataset, after binarizing the values of the pixels. As one can notice in Figure 3.5(a), the model considering the spiking micropillar shows better performances with respect a direct computation from the input - meaning that more information is encoded in each value of the binary output matrix X_{out} - in particular when the number of time bins is reduced. As represented in Figure 3.5(b), the performances of the software-based computation tend to overcome the ones from the model only for small training datasets.

Even though the values of the parameters used in the model are only qualitative, it is important to analyze the influence of hyperparameters on the system's behavior in order to gain a clear understanding of which inputs would be more suited for the experimental implementation of the algorithm. As represented in Figure 3.6(a), very low values of the bit time τ_b are reflected in a strong decrease in the accuracy. Indeed, many spikes cannot be triggered because of the refractory period of the system and the total amount of information given to the system is strongly reduced. However, the accuracy becomes approximately constant with $\tau_b \gtrsim 75$.
Figure 3.6(b) reports the dependence of the system from the threshold used for the binarization process during the decoding of the output data (or for the binarization of the input in the case of the direct computation).
Figure 3.6(c), finally, presents a heatmap of the average classification accuracy as a function of the parameters $\mu_{1,0}$ and c in Eq. 3.1. As expected, low values of $\mu_{1,0}$ correspond to a higher distance from the excitability threshold, resulting in a reduced number of spikes and, consequently, lower classification performance. As $\mu_{1,0}$ increases, this distance decreases, leading to a denser spiking response and improved accuracy, which eventually tends to saturate. The parameter c, which controls the strength of the input perturbation, plays a similar role in modulating the system's responsiveness. The saturation of the output may be interpreted as reduced sensitivity to these parameters over determined values, which could be useful during experiments.

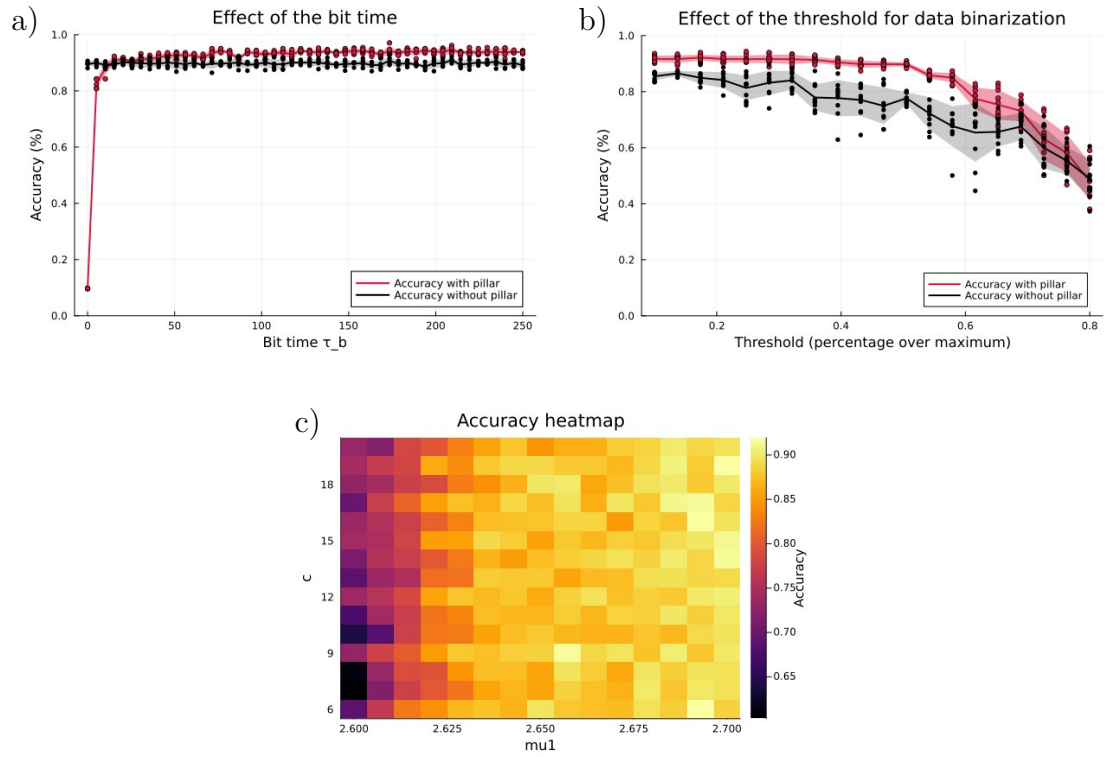

Figure 3.6: Accuracy dependence on the hyperparameters. (a) Dependence with respect to τ_b; $\mu_{1,0} = 2.68$, $c = 17.5$. (b) Dependence on the binarization threshold; $\mu_{1,0} = 2.68$, $\tau_b = 125.0$, $c = 17.5$. (c) Heatmap of accuracy vs $\mu_{1,0}$ and c; $\tau_b = 125.0$

3.3 Experimental analysis

3.3.1 Experimental setup

To validate the proposed approach, the algorithm outlined in Section 3.1 will next be deployed in a physical experimental setup, allowing performance assessment under realistic operating conditions and noise influence. A schematic representation of the experimental setup is shown in Figure 3.7.

The experimental apparatus contains two different continuous wave (CW) lasers and is divided into two main sections: the first employs fiber coupling to control the RF component of the pump, thereby generating the perturbations; the second operates in free space, where both the DC and RF components of the pump signal are delivered to the spiking micropillar.

The signal derived from the images is generated by modulating the output of a

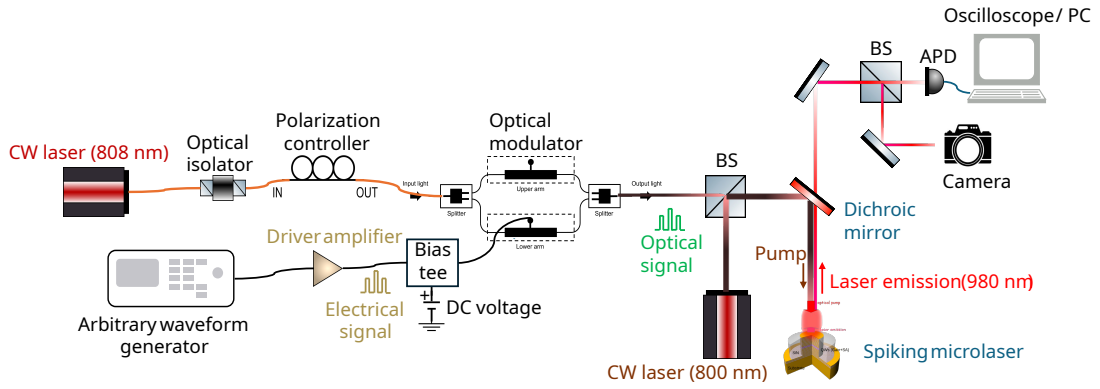

Figure 3.7: Schematic representation of the experimental setup

CW semiconductor laser (*Thorlabs MCLS1*), which will be denoted as L1, operating at a wavelength of 808 nm with a maximum optical power of approximately 30 mW. To minimize drift, the laser features internal temperature stabilization, maintaining a constant temperature of 25.00°C. The fiber-coupled output passes through an optical isolator to suppress any potential back-reflections, and is then routed through a polarization controller before being injected into an optical modulator (*iXblue NIR-MX800-LN-10-00-P-P-FA-FA*).

The waveform given as a RF voltage input to the modulator is created by an arbitrary waveform generator (*Tektronix AWG5002*), amplified by a *iXblue DR-VE-10-MO* driver amplifier, and combined with a DC voltage of $12\ V + V_{bias}$ via a bias tee. The role of V_{bias} is to effectively set the modulator at a specific point on its transfer function.

The optical signal that comes out of the modulator is then sent in free space to the second part of the experiment, where it is combined through a beam splitter with the light beam emitted by a second semiconductor CW laser (*Coherent FAP laser diode*), denoted as L2, with a wavelength of approximately 800 nm, which is responsible for the base (DC) component of the micropillar pump. The complete optical pump is then reflected by a dichroic mirror and directed to the spiking microlaser. To maintain optical performance, the micropillar is cooled and stabilized at a temperature slightly below 0°C using a Peltier heat pump. Exploiting the wavelength difference between the input and the output of the latter, the outcoming signal is able to pass through the dichroic mirror and to be directed to an RF-coupled avalanche photodetector (APD, *MenloSystems APD210*) with a bandwidth between 1-1600 MHz, from which it can be analyzed by a *LeCroy Wavemaster 8620A* oscilloscope with a sample rate up to 20 GS/s.

Working principle of the optical modulator

The *iXblue NIR-MX800-LN-10-00-P-P-FA-FA* optical modulator is based on a Mach-Zehnder interferometer, in which one arm incorporates a lithium niobate ($LiNbO_3$) crystal. Modulation arises from the Pockels electro-optic effect intrinsic to lithium niobate, whereby the application of an external voltage alters the crystal's refractive index. This voltage-induced index variation changes the optical path length, thus modifying the interference pattern at the output. The resulting intensity modulation is described by the following transfer function:

$$T(V) = \sin^2\left(\frac{\pi V'}{2V_\pi}\right) \tag{3.4}$$

with $V' = V - V_{off}$. In this expression, V_π denotes the voltage required to induce a full swing between the maximum and minimum transmission, while V_{off} is the bias voltage that results in a zero (or multiple of 2π optical phase difference between the interferometer arms). From a fitting procedure of the transfer function in Figure 3.8, the parameters were determined to be $V_\pi = 4.32\ V$ and $V_{off} = -2.07\ V$. A non-negligible thermal drift can be denoted in different sweeps of the curve.
Since this electro-optical response is essentially instantaneous - requiring no thermal response - and the device is engineered for high RF bandwidth, this system enables modulation with rise times in the order of 100 ps.

Possible working conditions

As represented in Figure 3.8, adjusting the peak voltage of the AWG directly determines the peak amplitude of the modulated waveform.

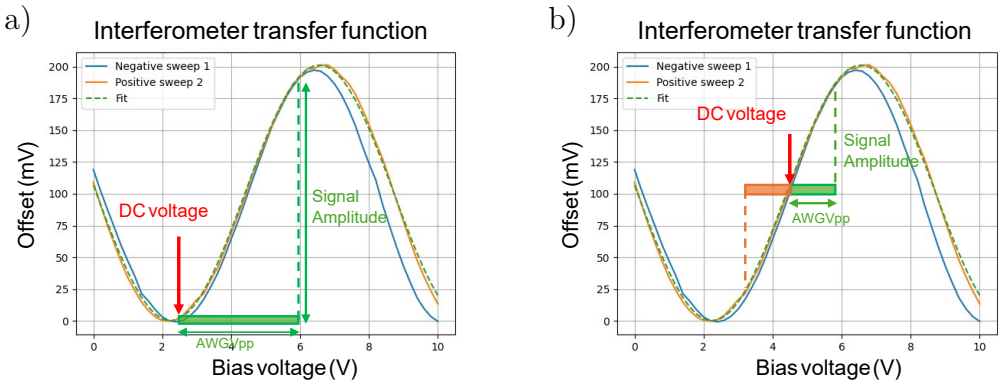

Figure 3.8: Possible working conditions of the optical modulator: (a) $V_{bias} = V_{off} + V_\pi$. (b) $V_{bias} = V_{off} + 3/2\ V_\pi$

Since we aim to encode pixel intensities into pump amplitudes, we can actively control the amplitude of the signal output from the optical modulator by adjusting the peak voltage of the waveform generated by the AWG. Meanwhile, the DC voltage determines the operating point of the modulator. One first possibility is to set the DC voltage V_{bias} in order to set the modulator in the minimum point of the transfer function in Eq. 3.4. This condition is achieved by applying a $V_{bias} = V_{off} + V_{\pi}$, as represented in Figure 3.8(a). This approach allows for a wider dynamic range in terms of the waveform amplitude.

The second possibility, illustrated in Figure 3.8(b), is to set $V_{bias} = V_{off} + 3/2\ V_{\pi}$ to work in the function region that may be approximated as linear: in this case the dynamical range will effectively be smaller, but it would allow the use of negative pulses, which are reflected in an effective pump removal.

Given the non-linearity of Eq. 3.4, it is useful to linearize the relation in order for the optical signal to maintain a coherent amplitude with respect to the one coming from the electrical signals that are applied by the AWG. To do so, the the voltage V in the transfer function can be computed through the inverse formula:

$$V'(x) = \frac{2V_{\pi}}{\pi} \arcsin\left(\sqrt{x}\right) \tag{3.5}$$

The modulated input is represented in Figure 3.9 in both the working conditions.

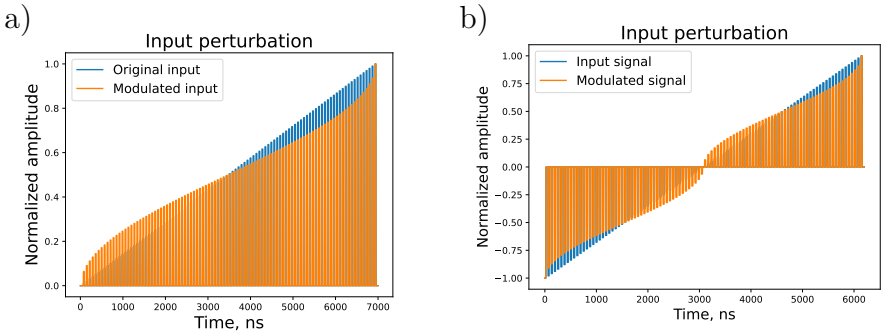

Figure 3.9: Linearization of the input from the AWG. (a) First working condition. (b) Second working condition.

Characterization of input signal

After modulating the input that is given to the AWG according to Eq. 3.5, the linearity of the input can be checked by applying a modulated perturbation to the AWG and recording the resulting signal with an oscilloscope. The results are reported in Figure 3.10.

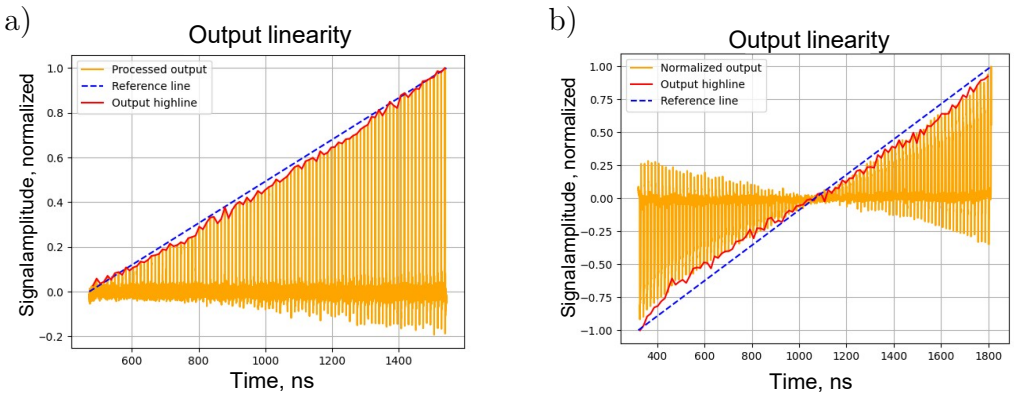

Figure 3.10: Check of the input linearity in the two working conditions.

3.3.2 Characterization of excitability

By adjusting the bias current of the laser L2 (I_{L2}), the bias voltage V_{bias} applied to the optical modulator, and the peak voltage V_{pp} of the waveforms sent to the AWG, a clear excitable behavior of the device can be observed. As illustrated in Figure 3.11, the emission intensity of the micropillar exhibits distinct spiking behavior when a perturbation is applied to the pump. To simplify the understanding, the two images are manually aligned using a marker (green waveform) as a reference.

As represented in Figure 3.12, the occurrence of spikes in the experiment is not perfectly deterministic. Due to the low power density of the pump perturbation laser on the sample — measured to be approximately 19 mW — it is necessary to set the bias current of laser L2 to values that bring the system very close to the excitable threshold. This operating condition results in an unstable output signal, where some spikes are triggered or suppressed by noise. The waveforms in Figure 3.12 are recorded using the oscilloscope in persisence mode, conserving a trace of each spikes for a time span of 2 s. The parameters used for Figure 3.11 and 3.12 are: $I_{L1} = 80.0\ mA$, $I_{L2} = 14.77\ A$, $V_{bias} = 2.32\ V$, $V_{pp} = 200\ mV$. Although the excitable behavior of the device is evident from the discrete nature of spike generation, the low optical power density of the perturbations results in a highly variable spike latency. This variability is evident by looking at the rapid decrease of spike latency (Figure 2.6) with an increase of the perturbation amplitude.

Figure 3.13, instead, reports the spiking activity of the sample when the input consists in a series of 20 pulsed perturbations with increasing amplitude. The expected all or none response is here clearly evident, as perturbations with a lower amplitude are not able to excite a spike; in contrast, larger perturbations elicit a spike with calibrated amplitude.

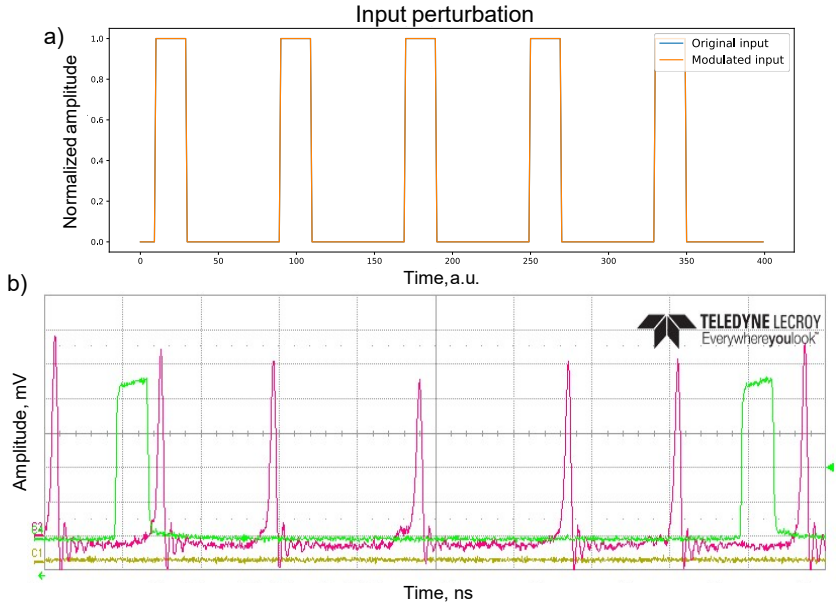

Figure 3.11: Experimental observation of excitability. (a) Perturbation scheme applied to the AWG, with high time $\tau_h = 4\ ns$ and low time $\tau_l = 10\ ns$. (b) Signal emitted by the spiking micropillar. The green waveform serves as a trigger for synchronization and signal acquisition. Each time division represents 10 ns, while each voltage division $20mV$.

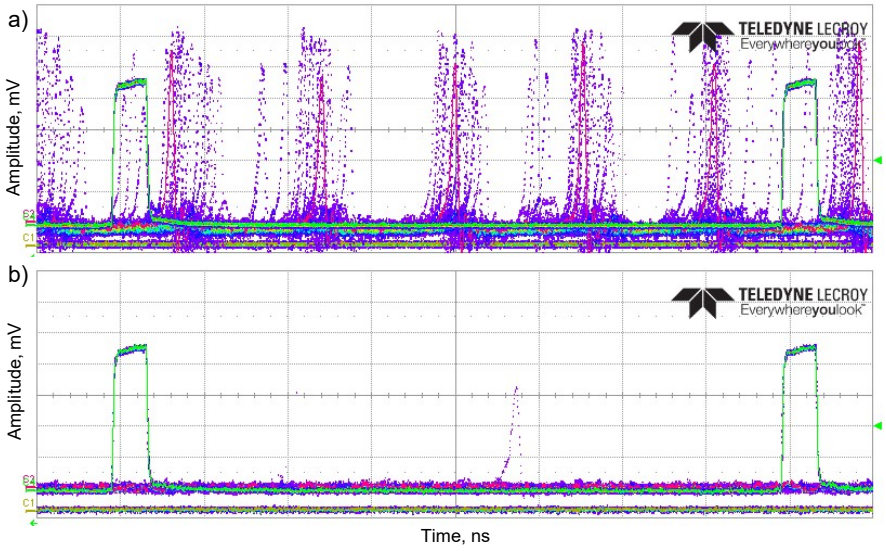

Figure 3.12: (a) Statistical temporal positioning of the spiking pattern. (b) Signal when no perturbation is applied to the pump. Each time division represents 10 ns, while each voltage division $20mV$.

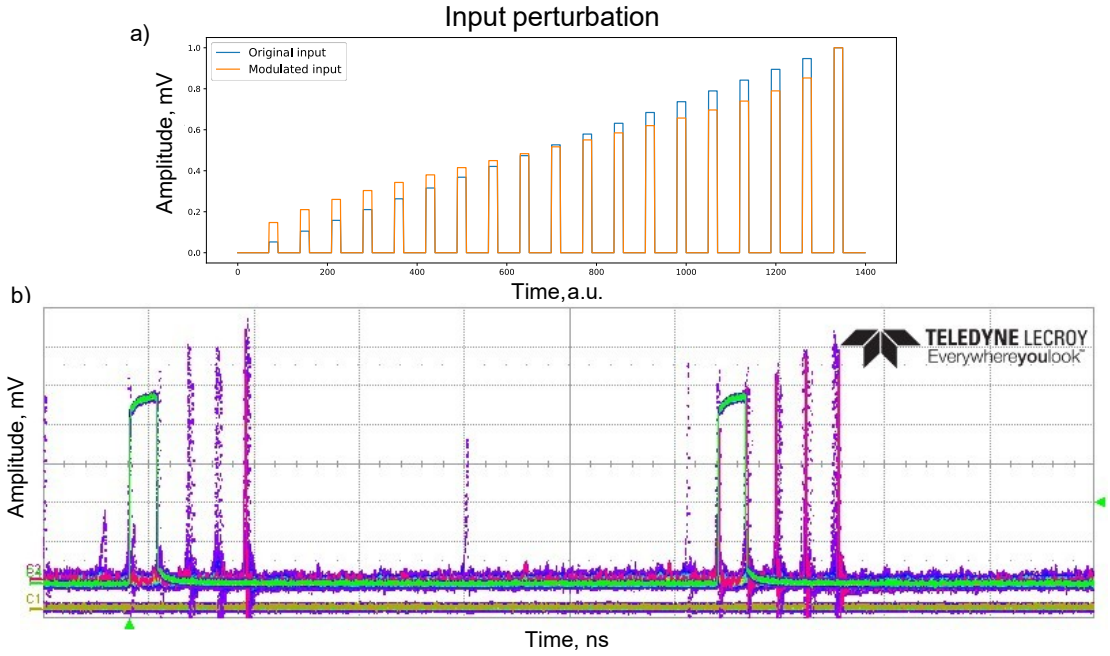

Figure 3.13: (a) Perturbation scheme applied to the AWG. (b) Signal emitted by the spiking micropillar; $\tau_h = 4;\text{ns}$, $\tau_l = 10\ ns$. Each time division represents 10 ns, while each voltage division $20mV$.

Chapter 4

Conclusions

This internship was a valuable opportunity to explore the research fields of photonics and to deepen my knowledge of neuromophic computing, a topic I have been interested in for a long time. During the course of the project, I became familiar with the physical principles and dynamical behavior of excitable micropillar lasers from theoretical, numerical, and experimental perspectives.

From a numerical standpoint, I implemented and analyzed vast simulations to characterize the properties of a single micropillar, and I managed to implement a complete image classification framework based on a reservoir computing architecture using an excitable microlaser as a single dynamical node, studying the influence of hyperparameters. By optimizing the simulation parameters and leveraging on the reading of each image from three different reading orders, I finally managed to obtain similar accuracy results tasks compared the PhD thesis this work was based upon [1] - which employed logistic regression as an optimization algorithm for the readout matrix - but employing a way less computationally expensive algorithm based on the Moore-Penrose pseudoinverse matrix. By exploiting Julia's multi-threading capabilities on an Intel i7 processor with 8 cores, the average computation time required to train the model with 125 samples per digit was approximately 50 s. This result clearly highlights the potential computational efficiency that a physical system such as the one described could help to achieve.

From an experimental point of view, I practiced the use of a complex optical setup. Although some limitations in the current configuration prevented full implementation of the classification task, the system's excitability was successfully demonstrated under controlled perturbations.

4.1 Perspectives

During the course of the internship some problems were found in the experimental setup, such as an insufficient power density of the pump perturbation onto the

sample. This is reflected in a clear non-deterministic nature of the spiking pattern, as one can notice in Figure 3.12. Because of this, and also due to time limitations, some work still has to be performed in order to produce a deterministic response in the spiking output to proceed with the experimental implementation of image classification tasks. In order to obtain a better control over the power density applied to the sample, a new more powerful laser has already been ordered by the group.

If better working conditions manage to be achieved, a single waveform containing the whole training dataset has to be created with a specific separation between each image. This waveform will be given as an input to the system and the results will be processed according to the algorithm described in section 3.1.1. An example of the perturbation scheme computed by a digit is reported in Figure 4.1

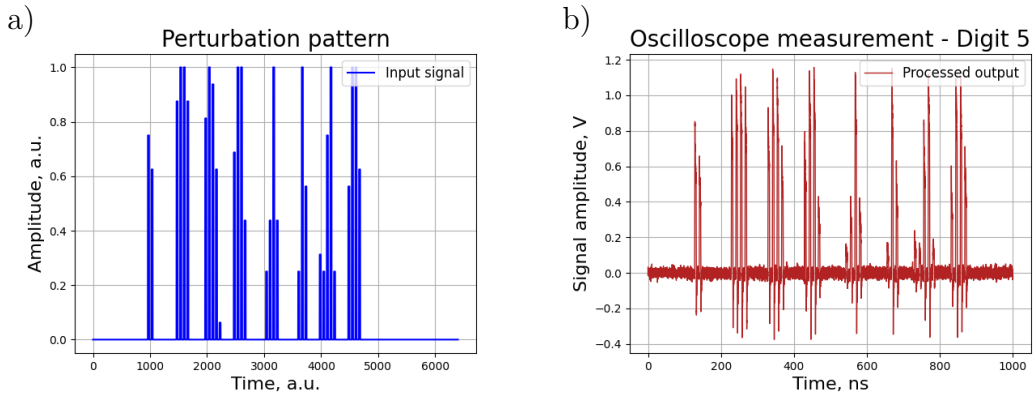

Figure 4.1: (a) Input applied to the AWG. (b) Signal after the optical modulator.

Recently, the group has been working also on a novel genetic algorithm training method [29] that could provide better performances and a high efficiency. A future perspective of this study could be to try to implement this into the physical system in order to improve its accuracy.
A final and ambitious goal, as anticipated in the introduction to this work, would be to create a fully optical ultrafast classification method by combining the results coming from different receptive fields with the use of on-chip photonic logic gates based on pulse collisions [18], on which I already deepened my knowledge by performing numerical simulations that were not reported in this report.

Bibliography

[1] Amir Hossein Masominia. «Neuro-inspired computing with excitable microlasers». Theses. Université Paris-Saclay, July 2024 (cit. on pp. 1, 3, 5, 9, 11, 13, 31).

[2] Gordon E Moore. «Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp. 114 ff.» In: *IEEE solid-state circuits society newsletter* 11.3 (2006), pp. 33–35 (cit. on p. 1).

[3] Chenxi Sun et al. «A review of designs and applications of echo state networks». In: *arXiv preprint arXiv:2012.02974* (2020) (cit. on p. 2).

[4] Junhyuk Woo et al. «Characterization of the neuronal and network dynamics of liquid state machines». In: *Physica A: Statistical Mechanics and Its Applications* 633 (2024), p. 129334 (cit. on p. 3).

[5] Daniel Brunner, Laurent Larger, and Miguel C. Soriano. *Nonlinear photonic dynamical systems for unconventional computing*. 2021. arXiv: `2107.08874 [cs.ET]` (cit. on p. 3).

[6] Kashu Yamazaki et al. «Spiking neural networks and their applications: A review». In: *Brain sciences* 12.7 (2022), p. 863 (cit. on p. 3).

[7] Wenzhe Guo et al. «Towards Efficient Neuromorphic Hardware: Unsupervised Adaptive Neuron Pruning». In: *Electronics* 9.7 (2020) (cit. on p. 3).

[8] L. Appeltant et al. «Information processing using a single dynamical node as complex system». In: *Nature Communications* 2.1 (2011), p. 468 (cit. on p. 4).

[9] Indranil Chakraborty et al. «Toward Fast Neural Computing using All-Photonic Phase Change Spiking Neurons». In: *Scientific Reports* 8.1 (2018), p. 12980 (cit. on p. 5).

[10] Yichen Shen et al. «Deep learning with coherent nanophotonic circuits». In: *Nature Photonics* 11.7 (2017), pp. 441–446 (cit. on p. 5).

[11] Zengguang Cheng et al. «On-chip photonic synapse». In: *Science Advances* 3.9 (2017), e1700160. eprint: `https://www.science.org/doi/pdf/10.1126/sciadv.1700160` (cit. on p. 5).

[12] Sylvain Barbay, Robert Kuszelewicz, and Alejandro M Yacomotti. «Excitability in a semiconductor laser with saturable absorber». In: *Optics letters* 36.23 (2011), pp. 4476–4478 (cit. on pp. 5, 6).

[13] Amir Masominia et al. «Online spike-based recognition of digits with ultrafast microlaser neurons». In: *Frontiers in Computational Neuroscience* 17 (2023), p. 1164472 (cit. on pp. 5, 13, 14).

[14] Bhavin J. Shastri et al. «Photonics for artificial intelligence and neuromorphic computing». In: *Nature Photonics* 15.2 (2021), pp. 102–114 (cit. on p. 6).

[15] Johan LA Dubbeldam, Bernd Krauskopf, and Daan Lenstra. «Excitability and coherence resonance in lasers with saturable absorber». In: *Physical Review E* 60.6 (1999), p. 6580 (cit. on p. 6).

[16] Miguel A Larotonda et al. «Experimental investigation on excitability in a laser with a saturable absorber». In: *Physical Review A* 65.3 (2002), p. 033812 (cit. on p. 6).

[17] Venkata Anirudh Pammi and Sylvain Barbay. «Micro-lasers for neuromorphic computing». In: *Photoniques* 104 (2020), pp. 26–29 (cit. on p. 6).

[18] L Soun et al. «Computing using pulse collisions in lattices of excitable microlasers». In: *Chaos, Solitons & Fractals* 164 (2022), p. 112537 (cit. on pp. 6, 32).

[19] T Elsass et al. «Fast manipulation of laser localized structures in a monolithic vertical cavity with saturable absorber». In: *Applied Physics B* 98 (2010), pp. 327–331 (cit. on p. 7).

[20] Johan L.A. Dubbeldam and Bernd Krauskopf. «Self-pulsations of lasers with saturable absorber: dynamics and bifurcations». In: *Optics Communications* 159.4 (1999), pp. 325–338 (cit. on p. 7).

[21] Minoru Yamada. «A theoretical analysis of self-sustained pulsation phenomena in narrow-stripe semiconductor lasers». In: *IEEE Journal of Quantum Electronics* 29.5 (1993), pp. 1330–1336 (cit. on p. 8).

[22] Pauli Virtanen et al. «SciPy 1.0: fundamental algorithms for scientific computing in Python». In: *Nature methods* 17.3 (2020), pp. 261–272 (cit. on p. 10).

[23] Venkata Anirudh Pammi et al. «Photonic computing with single and coupled spiking micropillar lasers». In: *IEEE Journal of Selected Topics in Quantum Electronics* 26.1 (2019), pp. 1–7 (cit. on p. 11).

[24] Fabian Pedregosa et al. «Scikit-learn: Machine learning in Python». In: *the Journal of machine Learning research* 12 (2011), pp. 2825–2830 (cit. on p. 16).

[25] Eric Hunsberger and Chris Eliasmith. «Training Spiking Deep Networks for Neuromorphic Hardware». In: (2016) (cit. on p. 19).

[26] Jun Haeng Lee, Tobi Delbruck, and Michael Pfeiffer. «Training Deep Spiking Neural Networks Using Backpropagation». In: *Frontiers in Neuroscience* Volume 10 - 2016 (2016) (cit. on p. 19).

[27] Chankyu Lee et al. «Deep Spiking Convolutional Neural Network Trained With Unsupervised Spike-Timing-Dependent Plasticity». In: *IEEE Transactions on Cognitive and Developmental Systems* 11.3 (2019), pp. 384–394 (cit. on p. 19).

[28] Wenjie Luo et al. *Understanding the Effective Receptive Field in Deep Convolutional Neural Networks*. 2017. arXiv: `1701.04128 [cs.CV]` (cit. on p. 21).

[29] Gibaek Kim et al. «Spike-based Classification with ultrafast Microlaser Neurons using Surrogate Model-assisted Genetic Algorithm training». July 2024 (cit. on p. 32).