Master 's Double Degree in Nanotechnologies for ICTs and Quantum Devices

POLITECNICO OF TORINO AND UNIVERSITÈ DE PARIS

Master Thesis

Nanophotonic gates for neuromorphic computing

Supervisors

Dr. Alfredo DE ROSSI

Prof. Maria Luisa DELLA ROCCA

Prof. Carlo RICCIARDI

Candidate

Luca AIMONE GIGGIO

June 2020

Acknowledgements

To my family. Without whom, I wouldn't be who I am.

First, I would like to thank Dr. Alfredo De Rossi, research member at Thales Research Technology - Paris, who accepted me for this internship and always managed to be available in spite of the smartworking condition in which much of the internship has taken place. He helped me not only in writing my report but also in dealing with all the issues related to this work.

My thanks goes also to Maxime Demulle for the great supervision and availability. He has in fact taught me, thanks to his previous work, a new approach to programming.

Foreword

The work that will be presented in this master thesis has been developed at the Thales Reserarch and Technology Physics group. The group consists of Dr Alfredo de Rossi and Dr Sylvian Combriè together with the two PhD students Maxime Demulle and Lèa Constants. In particular the group is focused on the development of all-optical processing gated based on fast nonlinear effects in order to control light using light. Especially the following work has been developed under the supervision of Dr Alfredo De Rossi together with the help of Maxime Demulle, who aim to use this technology to built neural network systems in collaboration with other research groups across Europe in the Fun-COMP project.

Table of Contents

List of Tables

List of Figures

Acronyms

ML
: machine learning

ANN
: artificial neural network

RNN
: recurrent neural network

FCD
: free carrier dispersion

NN
: neural network

PhC
: photonic crystal cavity

Chapter 1

Introduction

1.1 Context

This master thesis is placed within the modern context of the information processing where low-power hardware and software is needed for handling a large amount of data in the most efficient way. Here, the aim is to contribute to the recent interest in photonic technologies for increasing the computing power, specifically to address the big data challenge. This is motivated by the peculiar advantages of photonic technology in terms of speed and power efficiency.
The TRT research group where my internship experience took place aims to use integrated photonic devices in order to develop constructive elements able to mimic the essential characteristics of brain-like synapses and neurons in the interest of lay the foundations for low power and fast non-von Neumann computing. In fact, as powerful as might be an electronic chip is based on the 1940s John von Neumann architecture, this involves connections carrying data back and forth between memory and processor, and operates in strictly sequential way. Instead an artificial neural network is intrinsically parallel and decentralized. Therefore, the hope of neuromorphic computing is to overcome the bottleneck inherent the sequential operation of the von Neumann architecture.
The future objective of the group will be to develop integrated processors able to perform machine learning activities and therefore interface with real computational problems such as big data analysis.
This is the goal that Dr. De Rossi's team in collaboration with other researchers from all over Europe aims to achieve.

Having defined the context within which this work belongs, we will now define the objective of the following master thesis.

The photonic technology considered here is based on a nonlinear resonator which is operated as an artificial "neuron". In order to achieve energy-efficient operation, meaning use of little energy for each computing operation, ultra-small cavities are used, based on the technology of photonic crystals. These operate as an all-optical gate. I will study their characteristic starting from a theoretical model.
The specific point I have been addressing is the connection among a hierarchy of models,

from the most complex to the simplest (static), by analyzing the implications of the approximations made. The use of a lightweight but realistic model is essential in view of simulating the operation of a large network of such artificial neurons.

To this purpose, I will use particular Python platform called Photontorch, a library specifically defined to develop and model photonic elements and photonic neural networks under development by a group of researchers from the university of Gent [1]. It was perhaps the interface to this library the most difficult aspect of my work, in fact it is a high-level language developed to perform very specific tasks and operations. Moreover, as it is not yet commercially available, the documentation is not available.

1.1.1 Computing with light: a brief survey

Scientists started to dream about an optical computer from the 1960, and initially research focused on a digital transistor-equivalent device to control the on-off state of light. Researchers of the time claimed that such a machine could operate much faster than an electronic computer because of its higher light transmission speed and the possibility for parallel processing. Examples of the first optical implementations of artificial networks can be traced back in to the Hopfield network and the volume holographic neural network. The first neuronal models were in fact based on the concept of analog neuron defined through a non-linear activation function [1].
The Hopfield network, 1985, was based on the model for Hopfield neural networks and was also characterized by a nonlinear iterative feedback. The technology thus developed showed that the optical system was accurate and stable while maintaining all the advantages of optical technology, such as parallelism and the ability to stretch a large number of interconnections[2]. The machine had 32 neurons and had the capacity to store up to 1000 different values for interconnect weights. However, its performance was not comparable to that of the electronic systems of the time. For this reason a few years later, in 1988, a more complex version of the Hopfield network was developed: the volume holographic neural network. It consisted of a 3 dimensional implementation of the previous system, thus being able to store billions of weights[3].
Another example dates back to 1993, where an optical two layer neural network was used for facial recognition. The recognition accuracy was almost 100% and the it was instantaneous [4].
However, this technology did not prove to be competitive at time mainly for two reasons:

- these systems proved to be competitive only for large networks with extremely complex architecture
- the non-interaction characteristic of photons meant that they could not be used directly to control other photons, so switching required an indirect approach by modifying an intermediate material using electronics. However, the energy required for such a non-linearity was too high and therefore the development of an optical gate was considered impractical

[1]This type of artificial neuron will be defined in more detail in the paragraph 2.2.

which is why this technology was almost completely abandoned.
In addition electronics had been keeping on becoming smaller and faster. In recent years, however, Moore's law is beginning to waver, and the need is now to develop a technology capable of overcoming the limits of today's electronics[5].

Examples are the use of Mach Zender patterns to develop programmable nanophotonic processors [6], or the application of ring resonators to develop more advanced system based on a spiking neuron model [7]. These and many other approaches have been proposed in recent years generating an almost infinite range of possibilities.
To conclude, today, despite the enthusiasm of many optical ANN researchers, many scientists remain skeptical of their chances of becoming competitive with current digital systems. However, there is no doubt that Moore's law is currently in check and in the face of this reality photonic technologies have attracted significant interest.

Chapter 2

Non conventional computing

Below are the basic concepts relevant to contextualize the work of my master thesis. The reasons that led to the attempt to develop an unconventional method of calculation and how this can be contextualized within the modern reality of machine learning will be presented.

2.0.1 Towards parallel computation and problems

Von Neumann's architecture is characterized by a structure where CPU communicates with memory and operations are executed sequentially according to the Turing model. This sequential structure for data communication has several advantages, in particular it is much easier to develop algorithms able to work in this condition. However, in modern CPU, it was necessary to introduce a parallelism because the clock rate, at which operations are executed, could not be further increased without causing overheating. Parallelism has turned out convenient, in spite of a growing complexity. Yet, the communication between processors involves a bottleneck as it takes more than the 50% of the energy budget.
Faced with these problems, two main ways are possible to overcome the problem:

- The development of a non sequential computing paradigm (non Von Neumann) capable of a much better use of parallelism, in principle.
- The use of optical technology, motivated by their intrinsic capability to transmit information with low energy consumption.

2.0.2 Analog computation

Analog computation uses physical laws instead of algorithms to perform some calculations. In order to understand the difference between these two concepts a practical example can be made taking into consideration the Fourier transform. This operation can be carried out by a digital computer as a series of sum and subtraction with desired precision. For what concern the analog approach, if we consider a diffraction pattern, is known that the intensity profile distribution at the aperture is the Fourier transform of the diffraction pattern. The use of an analog calculation means that the accuracy depends on the correctness of the physical model used and also on the signal-to-noise ratio within

which the calculation takes place.
An example of analog computing implemented in a CMOS technology, is TrueNorth, a neural network developed using electronic circuits modeled as neurons [8].

2.1 Machine learning

Machine learning (ML) can be defined as the study of computer algorithms that improve automatically through experience[9], namely it consists in the development of an algorithm able to observe a quantity of data and define among them by means of a function (called decision function) a pattern through a series of of examples provided. This current of studies stems from the fact that computers today are not very adept at solving human-like tasks, and it is in this context that the machine learning research tries to define systems that are able to approach problems more generally, and learn from examples like humans. In order to achieve this goal, a particular technique is typically used within the field of ML, namely artificial neural networks (ANN). This tool is a computing system inspired by biological neural networks that constitute brain [10]. Starting from the neural network approach many subfield exist.
ANN have known a remarkable success as software implementation on a traditional von Neumann architecture. In the last years, the growing concern about rising carbon footprint of big-data has given a strong push to alternative approach based on hardware implementation of ANN, mainly using CMOS electronics, as TrueNorth[8], which however due to the high computer power consumption and the von Neumann architecture limitation are not very efficient.
More in general nowadays, many researchers are dedicated to the study of this new technology, in order to develop and characterize what are called neuromorphic devices. Namely a on-chip structure inspired by a neural network and with the same goal of solving problems now confined to the software field of machine learning. From a hardware point of view it is therefore completely different from the classical Turing machine architecture of modern computers.
Many advantages came from this new approach, in fact a non-Von Neumann computing eliminates the bottleneck between memory and processor, due to sequential structure between memory and CPU, by performing computation at the location of data. In this way one potential path towards lower power consumption and speed on specific tasks, such as the training of neural networks, can be achieved[11] .

2.2 Artificial Neural Network

An Artificial Neural Network (ANN) is basically a model which try to emulate the brain structure. It consists of "neurons" and connections, just like the brain, and in the same way recognize patterns in data by adjusting the strength of connections between the processing units (e.g neurons). In particular, depending on the needs of the model and the problem to be solved there are different types of neuronal elements which can be implemented.

2.2.1 Neuron Typologies

- **Threshold gates**: Is a category of neurons corresponding to a binary classifiers which only produces digital output. The main advantages of this class is that they are very simple neurons and universal for input-output digital computation.

- **Neurons based on an activation function (analog neurons)**: This class is based on computational units that evaluate the output through a nonlinear activation function and on a system of weights to manipulate the connections within the single neuron and between different neurons. In this way different output can be obtained according to the weighted sum of the inputs and the nonlinearity of the activation function.
  A widely used activation function is the sigmoid function, in fact, since is defined between 0 and 1, it useful for models where the output is predicted as probability. Another widely used function is the hyperbolic tangent, which compared to sigmoid has the advantage of being defined between -1 and 1, hence the negative inputs are mapped negatively and the zero inputs are mapped near zero in the tanh graph.
  An important characteristic of these neurons is that they support learning algorithms based on gradient descent technique [1].
  The input-output relationship of this kind of neuron, considered within a set of neurons, can be written hence as follow:

  $$y_j = f(\sum_{i \in S_j} W_{ij} y_i) \tag{2.1}$$

  where W_{ij} is connection weight between neurons, S_j a set of neurons describing the connections from neuron i to neuron j, and y_i the activation level of the neuron.
  This is just a brief introduction to this type of neuron that will be deepened later in section 3.

- **Spiking neurons**: These are the most biologically realistic but also the most complex neurons to model. Has been theoretically proven that Spiking Neural Networks can perform more complex operations than analog neurons [13].

2.2.2 Network Topologies

The key concept behind a neural network consists of the idea of neurons connected to each other by means of connections that can be described in matrix terms, called "weight matrices" W_{ij}. The purpose of weight matrices is to describe between which neurons there is a connection and the strength of it, it is in fact through the definition of these parameters that the network is defined. In particular, two classes of neural networks, distinguishable according to their connectivity, can be defined:

[1] The gradient descent technique is an optimization algorithm used to iteratively minimize the weights of the function in order to minimize the cost function, i.e. the function that defines how much the network results are correct compared with the desired output. For more details on ML techniques and theory [12].

- **Feedforward Neural Networks (FFNN)**: This topology is characterized by the absence of recurrent connections. The neurons are divided on different layers and the information is transmitted from the input layer to the output layer in a sequential way through the intermediate levels called hidden layers. However this unidirectionality of information has some limitations, in fact, although the network topology is simpler, it is not possible to store temporal information. For these supervised networks the best training rule is the error-backpropagation rule, which consists in iteratively minimizing a suitably defined "cost function" describing how close the network is to the desired behaviour, through the continuous readjustment of the connection weights.

- **Recurrent Neural Networks (RNN)**: These networks are characterized by feedback loops. This topology has several advantages, in fact unlike the previous topology it is now possible to introduce and contain time information in the system. This implies that the information contained by the single neuron is not only due to inputs, but also to previous states stored in the neuron and therefore depends on all previous inputs. However, there are not many functional learning rules, in fact the convergence of the output is fairly slow. A further problem is that, contrary to the solidity ensured by the previous topology, these networks are very sensitive to small variations of the initial parameters and therefore not stable.

A schematic comparison between the two types listed above can be seen in figure 2.1.

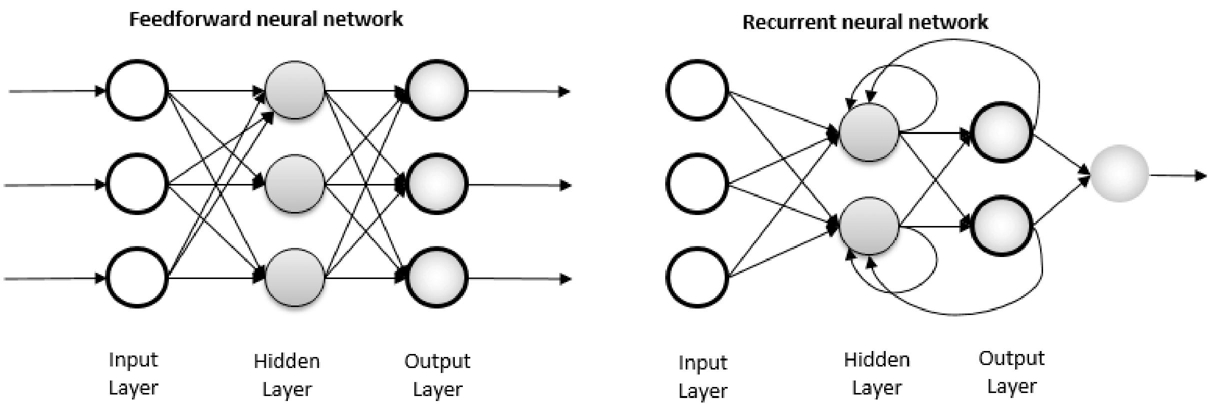


Figure 2.1: Comparison between Feedforward neural network and Recurrent neural network[14].

2.2.3 Reservoir Computing

This machine learning approach attempts to overcome the problems previously exposed related to training process in RNN. Different ways and networks have been proposed over the years, examples are: Echo State Network (ESN)[15], Liquid State Machine (LSM)[16], BackPropagation DeCorrelation (BPDC)[17].

In any case, the common idea behind this type of network, regardless of the proposed approach, is to transform the input signal to a high-dimensional state (e.g "reservoir")[2].

Initially RC was purely bound to software implementations to process temporal data, however, this approach has proved to be very versatile over the years, able in fact to deal with sequential problem (like speech recognition). Moving away from the software aspect in many ways has been tried to implement RC from a hardwer point of view [18]. Examples of this approach can also be found outside photonic technology like in memrisitive systems [19] and atomic switch network[20].

In general this type of network is based on a structure, shown in figure 2.2, consisting of three parts and characterized as follows:

1. An input layer whose job it is to pair the input signals.
2. The reservoir which represents the heart of the network and which in fact consists of an untrained recurrent neural network.
3. An output layer whose task is to linearly combine the states assumed in the reservoir in order to generate a time dependent output signal[21].

The procedure by which problems with RC are dealt with is typical of the supervised machine learning[3]. This means that training data are used to train the weights and get the output as a linear combination of the information stored in each node of the reservoir.

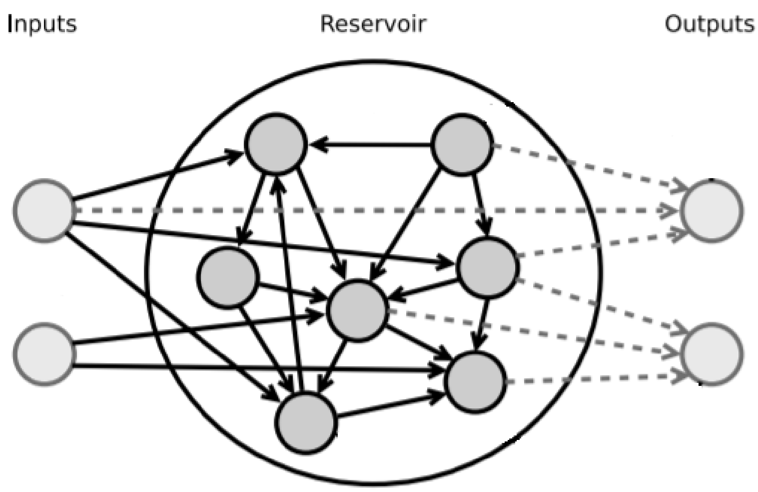


Figure 2.2: Graphic representation of a Reservoir Computing, consisting of a Recurrent Neural Network, an input layer and an output layer [22].

[2]This technique is called Kernel method, and is based on the fact that a non-linear surface is separable by a linear hyperplane in the high dimensional space. Therefore moving from a low dimensional input to an high dimensional space the performance of the system increases and the complexity of the problem decrease .

[3]More detail in 2.3

2.2.4 Training network

The main training techniques of this particular class of network will be quickly outlined below, which can be divided into two types: on-line and off-line

- **On-line learning**: The weights of the output are changed as soon as new information is transmitted in the network, this means that the learning procedure is done sequentially.
- **Off-line learning**: Using the dynamic responses due to the different inputs of the system, a comparison is processed with the desired output of the system in such a way as to change the output weights.

2.3 Learning Approaches

Before considering the physical properties of the neuromorphic device, will be discussed ML techniques, which are distinguished by how they implement learning in the network. In particular:

- **Supervised learning approach**: The goal is to build a function that maps an input to an output based on a known input-output pairs example [23]. Hence supervised learning algorithm analyzes the training data and produces an inferred function, which can be used for mapping new examples.
- **Unsupervised learning approach**: Is an approach that looks for previously undetected patterns in a data set with no pre-existing labels and with a minimum of human supervision. In contrast to supervised learning that usually makes use of human-labeled data, unsupervised learning, also known as self-organization allows for modeling of probability densities over inputs[24].
- **Reinforcement learning approach**: Is an area of machine learning concerned with how software agents ought to take actions in an environment in order to maximize the notion of cumulative reward. Reinforcement learning differs from supervised learning in not needing labelled input/output pairs be presented, and in not needing sub-optimal actions to be explicitly corrected. Instead the focus is on finding a balance between exploration (of uncharted territory) and exploitation (of current knowledge)[25].

2.4 Photonics

As mentioned above, one of the main issues with software implementation of ANN is the large amount of energy required. In general, the trouble with modern computer microprocessor is the energy cost transferring data within chip, which represents more than the 50% of the total power consumption and the situation deteriorates as the bandwidth and the complexity (i.e. number of gates) increases in the digital CMOS technology.

One step further is taken when considering moving from electrons to photons to implement artificial neural network. The very use of light as a means of transporting

information introduces in the system more degrees of freedom, in fact light is characterized by the amplitude and phase. Moreover the speed of the communication it is not limited by circuit element and is independent by the energy consumption. Other advantages arise from the fact that using photonic technology a large bandwidth for the signal is allowed and the losses of the system are almost negligible.
However some drawbacks are present, in fact the constructive tolerance in this photonic device must be taken into account since slight variations in the construction characteristics of each element produce large variations in the expected behavior of the photonic device and therefore the expected response from the system.

2.4.1 Computing primitive for Optical Neural Networks

As introduced above a good platform for the creation of artificial neural network is the photonic technology, in fact an artificial neural network, for instance an RC network (figure 2.2), basically consists of a large number of nonlinear elements connected to each other.
A nonlinear activation function is the essential element of a neural network. In order to implement such function, a varierty of technologies have been considered in the past years, like Semiconductor Optical Amplifiers (SOAs) and nonlinear ring resonators[18].
Here we consider photonic crystal cavities as promising for neuromorphic devices:

- Since they are passive devices, unlike SOA, they require less energy.
- The ability to store energy in very small volume enhances non linear effects (FCD, Kerr, temperature), as will be discussed in more detail later in 3.4.
- Photonic crystals are intrinsically bidirectional devices since it is possible to define a transmitted and a reflected signal, allowing the introduction of a possible feedback system.
- An additional degree of freedom in the network can be achieved through construction geometry. In fact the crystal cavity can be placed inside the guiding structure, namely "inline cavity", or alongside the waveguide, namely "side coupled cavity". The main difference between these two construction solutions is that the first is characterized by maximum transmission when the cavity is at resonance, while the second has zero transmission at resonance.
- The scalability of the cavity allows many of these elements to be placed on very small footprint.

Chapter 3

Model for a neuron

So far, the key concepts related to what is meant by neural network and how photonic technology can intervene in the development of such systems have been presented globally. In the following the discussion will focus in particular on the study and formalization of the basic element constituting such networks, i.e. an artificial neuron. A biological neuron is mainly composed of three elements: soma, axion and dendrites. The soma is characterized by a difference in potential and as a result of signals coming from outside, if a certain threshold of potential is exceeded, there is the emission of a nervous signal through the axon and dendrites to adjacent neurons[26]. This behaviour can be modelled through a system composed of a computing unit connected to other computing units by means of weights characterising the connection and a unit defining the activation function, which in biologically inspired neural networks is an abstraction representing the rate of action potential firing in the cell[27] beyond which there is the emission of the signal. A graphic representation is shown in the figure 3.1
As previously mentioned in the paragraph 2.2.1 this is therefore a model based on analog neurons.

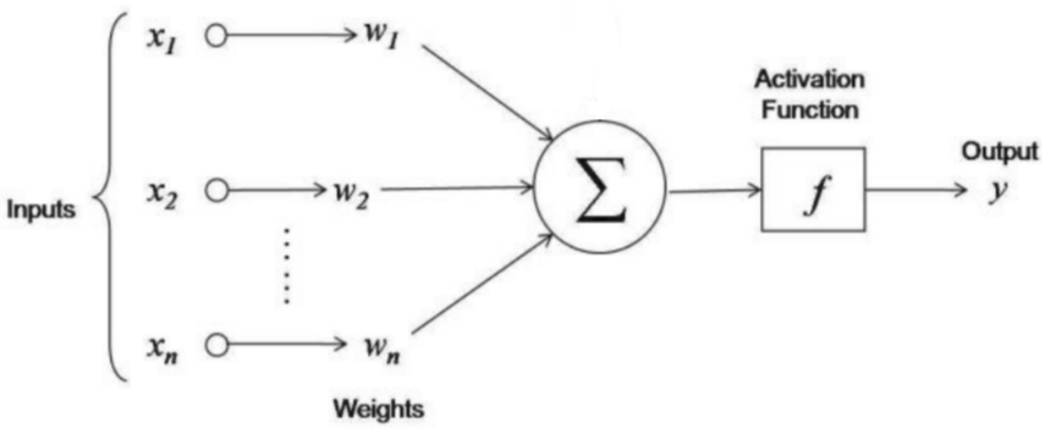


Figure 3.1: Model formalisation of an artificial neuron.

As hinted in the paragraph 2.4.1 the artificial neuron considered here is therefore a photonic crystal cavity. However, of fundamental importance for the characterization of an artificial neuron is the implementation of its activation function. In fact this represents the logical unit able to decide whether outside connections should consider this neuron as "activated" or not.
In the next paragraphs a static model describing a photonic crystal cavity will be examined to define its non-linear output characteristic (namely the artificial neuron activation function). The model described will start from a more complex dynamical treatment to which various simplifications will be made. The purpose will be then to analyze the limit of validity of the developed static treatment. As far as the theoretical part is concerned, the model that will be presented derives from a discussion still under development by Dr De Rossi, while the comparison algorithm that will be used for the dynamical model has been developed by Demulle.

3.1 Photonic Crystal Resonators

Photonic crystals are optical structures characterized by the possibility of modulating their properties periodically at wavelength scale. As a matter of fact, they can be interpreted as Bragg reflector, in fact like these devices, the coherently scattered waves add up in order to produce a reflection characterized by a strength that grows with the number of periods of the structure. This implies that a periodic structure does not allow the propagation of light in a certain spectral range, and therefore the photonic gap introduces in the system fundamental properties for the control of light matter interaction. The introduction of a defect in the periodicity of the structure creates a resonant localized states and in this way the photonic crystal resonators is achieved. Of fundamental importance for the resonator is therefore the manufacturing of the defect, on which the quality of the light confinement within it and the quality factor of the structure depend. However, even if today's manufacturing techniques are very advanced, it remains a challenging aspect in the resonator fabrication due to the tight tolerances required [28].
In particular, as mentioned in 2.4.1, there are two possible configurations for the location of the defect : the cavity can be located inside the guiding structure, "inline cavity",or outside the guiding structure, "side-coupled cavity". A schematic representation is shown in the figure 3.2

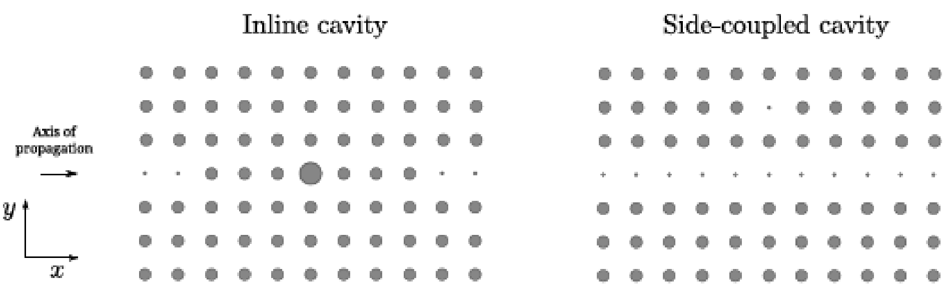


Figure 3.2: Possible architecture layouts[29].

3.1.1 2D Photonic Crystal

Theoretically, in order to confine light in all directions, it would be necessary to construct a three dimensional periodic structure, but this is not an easy manufacturing procedure. For this reason through a waveguide the light is totally confined and guided along the z-axis exploiting the total internal reflection, while along the (x,y) plane the properties deriving from the periodicity of the photonic crystal are used. In this way it is possible to obtain a confinement of light along the three spatial directions with a 2D PhC combined with a waveguide, as shown in figure 3.3.

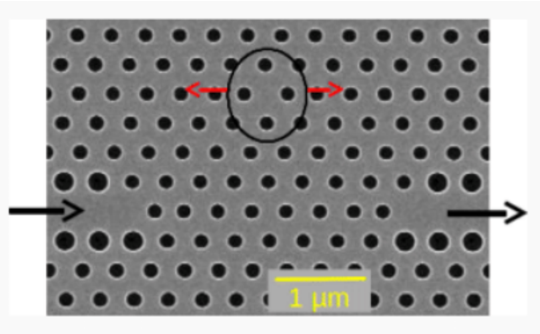


Figure 3.3: Side coupled PhC[30].

Considering a two dimensional photonic crystal with a lattice defined by two perpendicular vectors $\vec{a_1}$ and $\vec{a_2}$, having same magnitude a, in the (x,y) plane, it can be assumed invariant along z direction. The periodicity of the structure allows, as mentioned in the paragraph 3.1, the formation of a photonic band gap where propagating states are not allowed. In order to describe this behaviour a Bloch formalism approach can be applied.

In fact the system is characterized by a discrete periodic symmetry for the dielectric unit $\epsilon(r) = \epsilon(r + R)$ for any R integer multiple of any linear combination of the primitive vectors a$\hat{x}$ and $a\hat{y}$.

Applying the Bloch's theorem, it is possible to write the magnetic field as a function of the Bloch state and focus the treatment on the values of $k_{||}$ inside the Brillouen zone:

$$H_{n,k_z,k_{||}}(r) = e^{ik_{||}\rho} e^{ik_z z} u_{n,k_z,k_{||}}(\rho) \tag{3.1}$$

where n represents the band number, $k = k_z + k_{||}$ the wave vector, ρ the projection of r in the (x,y) plane and $u(\rho)$ the periodic function, namely $u(\rho) = u(\rho + R)$.
Due to the symmetry in the (x,y) plane and the invariance along z, the modes can be classified according to the their polarizations: respectively transverse electric (TE) modes and transverse magnetic (TM) modes.
The subdivision between these two modes has important consequences, in fact the photonic band structure depending on the mode can be completely different.
In these structures it has been demonstrated that for cavities with dimensions comparable to the incident wavelength, they are characterized by very strong light confinement. The quality factor (Q) for modal volume (V), Q/V, determines the strength of the interaction of the cavity with the mode, and for these devices is characterized by very high values.

However, realizing this type of structures is still sufficiently complex[31]. This is the reason why today nanophotonics is a field of research that attracts a lot of interest, in fact, considering the case of photonic crystal cavity, it allows to store the same amount of energy but in a much smaller volume. This implies that the nonlinear response from the cavity will be much greater since the properties of the material depend on the density of the electric field contained within it. This concept will be discussed in more detail in 3.5.1.

Next we introduce a dynamical model of a semiconductor PhC resonator [32]. It is described by dynamical variables corresponding to the field and to the population of free carriers excited by absorption in the material. Starting from this complex model several assumptions will be made to simplify it until it is completely developed in static approximation.

3.2 Coupled Mode Theory

Considering a system composed by a single resonator coupled with a waveguide and assuming per hypothesis the case of single mode resonator (figure 3.4a), the coupled mode theory describes the evolution of the field inside the cavity in terms of the evolution of its normal modes [33]. The field in the resonator is dependent to the amplitude A as:

$$E(x,t) = \sum_n A_n(t) u_n(x) e^{i\omega_{ref} t} \tag{3.2}$$

where ω_{ref} represents the reference angular frequency namely, the working frequency of the laser. The amplitude evolves in time according to the following motion equation:

$$\dot{A} = M \cdot A - R \cdot S_i \tag{3.3}$$

where it is possible to recognize the input electric field amplitude S_i, the term representing the effects of interaction with matter M, and a parameters depending on the geometry, R. In particular $M = i\Delta\omega - \frac{\gamma}{2} - \frac{\Gamma_0}{2}$ and $R = k$. Where $\Delta\omega = \omega_0 - \omega_{ref}$ is the cavity detuning from the laser, Γ_0 denotes the internal loss and γ the coupling to the waveguide.

3.3 Free Carriers

Absorption in semiconductors occurs, of course, when an absorbed photon allows the transition of an electron into the conduction band, leaving a hole in the valence band. These populations of free carrier can be represented by the density N(x,t), which is assumed under the hypothesis of ambipolar approximation to be the same for electrons and holes. The dynamical model describing the time evolution of the density can be written as follow:

$$\dot{N} = -\Gamma_c N + \left(\frac{1}{2}\Gamma_{TPA}|A|^4 + \Gamma_A|A|^2\right)\frac{1}{V_c \hbar\omega} \tag{3.4}$$

In order to simplify the equation the two photon absorption process (TPA) is neglected.

The simplified model can be rewritten as follow:

$$\dot{N} = -\Gamma_C N + \Gamma_A |A|^2 \frac{1}{V_c \hbar\omega} \tag{3.5}$$

where Γ_C and Γ_A are respectively the recombination rate and the absorption rate, ω the resonance frequency of the cold cavity and V_c the modal volume contained in the cavity.

Assuming in eq. 3.5 the recombination rate sufficiently large the carriers variation in time can be considered negligible since it is possible to regard the instantaneous rearrangement of these. Hence through this consideration it is possible to relate the carriers density distribution to the internal energy mode stored by the cavity:

$$N = \frac{\Gamma_A}{\Gamma_c V_c \hbar\omega} |A|^2 \tag{3.6}$$

3.4 Nonlinear response in semiconductors

In dielectric structures several non nonlinear effects can arise, typically they can be negligible, however in a cavity the intensity of the incoming field is strongly enhanced and then the nonlinear effects become relevant.
In semiconductors the dominant nonlinear effects are related to Kerr nonlinearity, and to the refractive index dependence on the temperature or the density of free carriers. Even though Kerr effect is much faster, free carrier nonlinear response effects can be stronger. In fact, is well known that the non linear index change strongly according to the concentration of the free carrier dispersion. The strength of the effect also depends on how close to the edge of the electronic bandgap, in fact when the photon energy is close to the bandgap, the non linear effect due to bandfilling appears and dominates the response.
Is proven that the classical free carrier plasma dispersion (FCD) effect described by the Drude model dominates when the plasma energy is well below the band gap and this condition is here taken into account.
In any case, the refractive index decreases while increasing the free carrier population due to FCD, inducing usually a blue shift in the cavity resonance. In particular it is experimentally demonstrated that for a photonic crystal the FCD induces a non linear response of few orders of magnitude larger than Kerr effect [34].
Another important parameter is the effect of temperature. In fact, the refractive index may change due to the thermal refractive effects as dissipation increases the temperature, generating a red shift of the resonance.
Therefore, the two dominants nonlinear terms to take into account are free carriers and temperature. Then considering these two effects the change of the refractive index resulting into a change of the frequency of the resonance through:

$$\Delta\omega_{NL} = \partial_{N\omega} N + \partial_{T\omega} T \tag{3.7}$$

Where $\partial_{N\omega} N$ and $\partial_{T\omega} T$ represent the linearized dependence on T and N. However in the following discussion for simplicity the thermal effect will not be considered.

Including the nonlinear terms, the motion equation shown in eq. 3.3 becomes:

$$\dot{A} = MA - RS_i + i\Delta\omega_{NL}A \tag{3.8}$$

3.5 Static limit

Now considering a further approximation: the static condition for the motions equation $\dot{A} = 0$. Hence using eq. 3.6 and 3.7 in eq. 3.8 under steady state condition it is possible to write:

$$\left[i\left(\Delta\omega + \frac{\Gamma_A}{\Gamma_c V_c \hbar\omega}|A|^2\right) - \frac{\gamma}{2} - \frac{\Gamma_0}{2}\right]A = kS_i \tag{3.9}$$

A further approximation can be done, in fact in order to simplify the model, the internal loss Γ_0 is assumed zero. In this way the total linear energy damping $\Gamma = \Gamma_0 + \gamma$ can be approximated directly with the coupling term (e.g γ). Furthermore the coefficient k is determined by the conditions of conservation energy [35], hence in absence of internal loss can be defined $k = \sqrt{\gamma/2}$. It also introduces a parameter characterizing non linear effects, g, defined as follow : $g = \frac{\partial_{\omega N}\Gamma_A}{\Gamma_c V_c \hbar\omega}$ which under typical operating conditions is a positive value.
With these considerations the previous equation can be rewritten in the following way:

$$\left[i\left(\Delta\omega + g|A|^2\right) - \frac{\Gamma}{2}\right]A = \sqrt{\frac{\gamma}{2}}S_i \tag{3.10}$$

Applying then the modulus square to eq. 3.10 a cubic polynomial equation in $|A|^2$ is obtained, where the energy stored in the cavity (e.g $|A|^2$) can be reached according to the input power in the cavity (e.g $|S_i|^2$). Namely:

$$g^2|A|^6 + 2\Delta\omega g|A|^4 + \left(\Delta\omega^2 + \frac{\Gamma^2}{4}\right)|A|^2 - \frac{\gamma}{2}|S_i|^2 = 0 \tag{3.11}$$

The solutions for the cavity energy of the previous equation allow to solve eq. 3.10 from which the amplitude A of the mode stored in the cavity can be obtained.

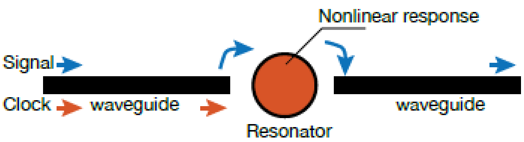
Depending on the configuration of the structure, namely in line cavity and side coupled cavity, the relationship between input and output fields takes a different form, respectively $S_o = \sqrt{\frac{\gamma}{2}}A$ and $S_o = \sqrt{\frac{\gamma}{2}}A + S_i$. In this discussion and in particular in the following chapters a in line configuration will be taken into account. Hence through the relationship written above it is possible to obtain the output as a function of the input, from which a non linear behaviour is expected. But before analyzing the solutions of the model it is necessary to introduce and better define the phenomenon responsible for the expected non-linear feature, the so called optical bistability.

3.5.1 Optical Bistability and logical applications

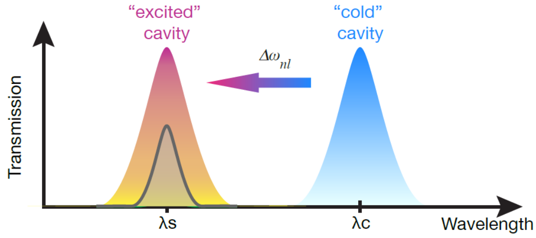
Taking into account in general a resonant nonlinear cavity coupled to an input and an output in linear regime and considering as input and output two single mode waveguides, figure 3.4a, the ratio between input and output powers is characterized by the typical Lorentzian shape:

$$\frac{P_o}{P_i} = \frac{1}{1 + (\frac{\omega_{ref} - \omega}{\Gamma})^2} \tag{3.12}$$

where Γ represents the total dissipation rate.



(a) In line cavity configuration.[28]



(b) Resonance shift under the effect of the detuning.[28]

Figure 3.4: Principle of resonant all-optical switch.[28]

Considering now the photonic crystal resonator with the nonlinear mechanism exposed previously, a non linear bistable response of the resonator is now taken into account and described.

Historically, bistability has represented an important step in the definition of an optical transistor and therefore to exploit this response in order to realize an optical gate equivalent to a CMOS gate. Today this approach is outdated because in today's neural network applications bistability is no longer used. However, it remains interesting to characterize the bistable operation to understand this technology [36].

Looking at the eq. 3.10 it is possible to distinguish two important parameters to characterize a bistable behavior: the detuning, $\Delta\omega$, and the effective detuning $(\Delta\omega + g|A|^2)$. The role of detuning is to set the frequency of the incident field outside the cavity resonance. Important for the occurrence of a bistable event is to set the detuning so as to have the opposite sign to the non-linear term g, in this way, as the energy contained in the cavity increases (i.e. as the input power increases) the non-linear phenomena will lead to the compensation of the detuning. In this compensation situation the effective detuning will be equal to zero, thus causing a shift of the resonance towards the operating frequency of the input signal, figure 3.4b. The system, therefore, entering resonance, will pass to another stable solution.

What happens, therefore, is that the system passes from an out-of-resonance state to a resonance state, thus presenting a transition between two stable states. The system is then called bistable because it presents two output states for the same input at the moment in which the transition takes place . A hysteresis cycle appears in the input-output plane as the input is varied cyclically. This is the hallmark of bistabiity.

Through this system an optical gate can be realized. A logical function can be associated to the two stable states of the system and thanks to the detuning control it is possible to dynamically control the gate. The advantage of the use of photonic crystals in this context lies in the efficiency they bring in terms of switching energy and switching time, in fact from comparisons with other photonic devices a good trade off between these parameters appears, as shown in figure 3.5. It is particularly by moving into the domain of nanophotonics that these advantages become fundamental. In fact the non-linear response are properties of the material and depend on the energy density contained locally. Therefore, by developing nanocavity it is possible to concentrate the same amount of energy in a very reduced volume.

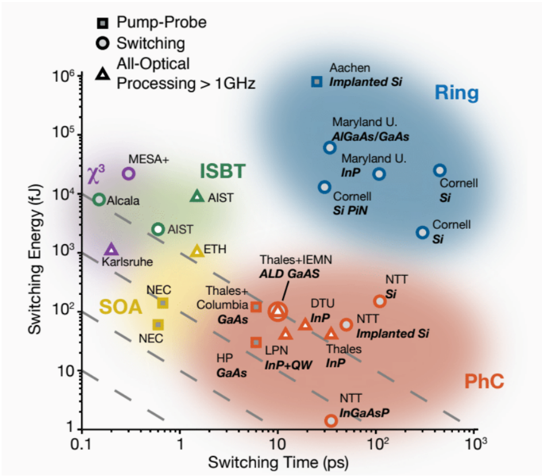


Figure 3.5: All optical processing compared[34].

The large and fast nonlinear response and the small footprint are of fundamental importance for implementing nonlinear activation function in neuromorphic devices as they provide new features with new logical functions that can be introduced in training systems.

Chapter 4

Towards a photonic NN based on PhC resonators

It is important to underline that the approach for the development of NN proposed here is substantially different from what can be found in similar research. Here in fact all the data used for the formulation and implementation of the model derive from the characterization of photonic crystals. Therefore, the model developed is based on an existing device whose characteristics have also been experimentally verified [34]. This is a big difference compared to other researches developed on abstract and not well defined elements. The parameters used and obtained experimentally are given in Appendix A.
Starting from this key consideration, from the model described in paragraphs 3.2-3.3-3.4, we will now define the methods and approach with which this has been addressed in the following discussion.
The physical model presented was approached using Python programming language, In particular it has been implemented on a particular library named "Photontorch".
Photontorch is a library developed by a group of researchers on the subject of photonic neuromorphic devices from the University of Gent. Dr Peter Bienstman has kindly granted the use of this platform for writing of this report, in the framework of FunComp H2020 project.
Photontorch is a tool for highly parallel simulation and optimization of photonic circuits in time and frequency domain. Photontorch features CUDA enabled simulation and optimization of photonic circuits. It leverages the deep learning framework PyTorch to view the photonic circuit as essentially a sparsely connected recurrent neural network. This enables the use of native PyTorch optimizers to optimize the (physical) parameters of the circuit [1].

4.1 Solving process

One of the major issues I faced in implementing our model was to learn how to use the Photontorch working environment. It is in fact a library still under development with a completely different approach compared to classic programming languages.
In order to implement out the "neuron" model with Photontorch was considered the cavity

as a system defined by 3 variables (e.g. ports): the input, the mode energy contained in the cavity and the output. $|A|^2$ has in fact been used as an internal variable of the system in order to keep track of the variations in the cavity response. To do this the model has been manipulated (see 3.4) in order to obtain a third degree polynomial equation in $|A|^2$, eq. 3.11. Once found the solutions of this equation only the one characterized by the value corresponding to the minimization of the energy contained in the cavity is accepted and stored as an internal variable of the system. Iteratively this process is implemented considering at each iteration as minimum energy state the one stored at the previous iteration.
So the solutions thus obtained were used within eq. 3.10 to calculate the amplitude of the mode within the 2D PhC. From these, finally, the amplitude of the response is evaluated through the relation previously exposed in 3.4 , namely $S_o = \sqrt{\frac{\gamma}{2}}A$ for in line cavity and $S_o = \sqrt{\frac{\gamma}{2}}A + S_i$ for side coupled cavity.
It is important to underline that in the following discussion all the values and results have been obtained by considering an in line cavity configuration.

4.1.1 Dynamical Model

Since the core of this work is to define a static model and to verify its convergence with the more general dynamical model for the photonic crystal cavity, it will be briefly defined below.
The dynamic model for the cavity with which the comparison will be made was developed by doctoral student Maxime Demulle in the Photontorch programming environment. The latter has been built according to the same assumption and parameters [1] as the static one, however, solving at an integral calculation level the amplitude of the electric field inside the cavity eq. 3.10.

4.2 Results and Discussion

In the following section the results obtained from the static model will be analyzed. In order to do this, the signal transmitted by the cavity will be studied. The treatment will be performed by dividing the analysis into these 3 points:

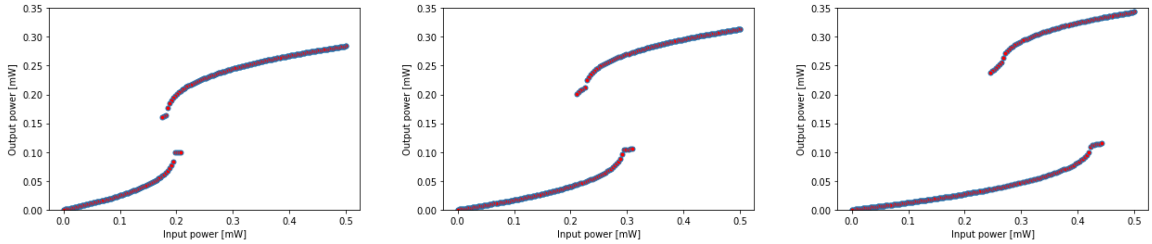
1. Hysteresis loop demonstration.
2. Detuning.
3. Limits of convergence between static and dynamical model.

4.2.1 Hysteresis loop demonstration

Using an increasing and decreasing ramp input power between 0.01 and 0.5 mW the power transmitted by the system is analyzed; what is expected is the appearance of a cycle of

[1]The parameters used are reported in the appendix A, table A

hysteresis, figure 4.1



(a) $\Delta\omega = -0.1^{12} rad/s$ **(b)** $\Delta\omega = -0.12^{12} rad/s$ **(c)** $\Delta\omega = -0.14^{12} rad/s$

Figure 4.1: Hysteresis loop for variable detuning.

As can be seen using this range of powers a hysteresis cycle appears, and in particular as the detuning value increases in magnitude the hysteresis cycle also becomes wider. The following phenomenon can be described by analysing separately the cavity's response to the increase and decrease in input power:

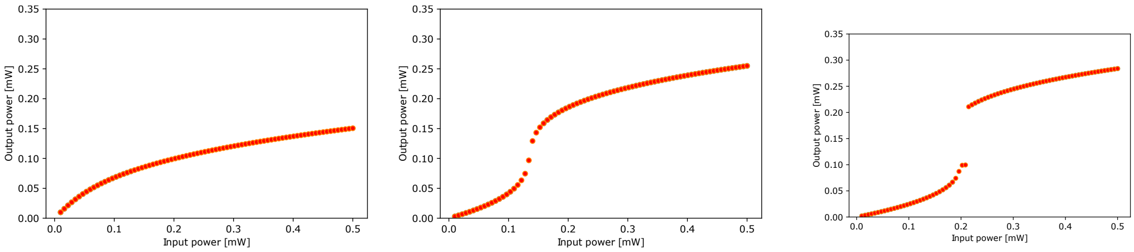
- **Increasing input power**: in the ascending range part [0.01:0.5]mW the system starts with zero input energy, which means that the initially stable solution of the system will be the lower one. This is until for sufficiently high energies the system moves to the upper solution when the resonance shift occurs.
- **Decreasing input power**: considering now the decreasing part of the input signal (0.5:0.01)mW, the system will be permanently in the upper solution. Therefore, starting from this solution, the system will pass through the stable states of this configuration until it transitions specularly with respect to the previous case in the lower solution.

The results reported here were obtained not through Photontorch, but through the same model developed on Python. This is because through the code developed on Photontorch an error for the moment unsolved appeared making it impossible to display the hysteresis cycle.

4.2.2 Detuning

As mentioned in the previous section, a very important parameter is the angular frequency detuning. In fact the cavity exhibits monostable or bistable behavior depending on the operating angular frequency ω_{ref}. In particular it's defined the angular frequency detuning as $\Delta\omega = \omega - \omega_{ref}$. Here we will study its effect on the static characteristic given at the input a constant signal between 0.01 and 0.5 mW.

What is shown in Figure 4.2 is detuning compensation through the non-linearities of the system.
In fact the effect of detuning is to modulate the external pulse in order to meet the shifting resonance from non-linear effects, as explained in 3.5.1. This involves the definition of two stable solutions: one out of resonance (lower part of the curve) and one in resonance (upper part of the curve), corresponding respectively to low transmission condition and



(a) $\Delta\omega$= 0 rad/s **(b)** $\Delta\omega$= - 0.08^{12} rad/s **(c)** $\Delta\omega$= -0.1^{12} rad/s

Figure 4.2: Non linear response of inline cavities for several detuning.

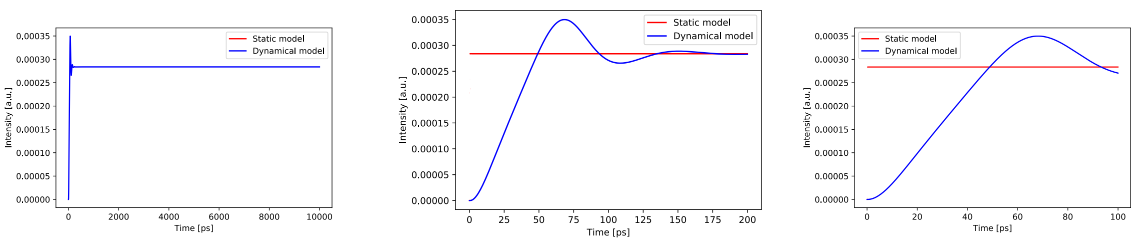
high transmission condition. The trends shown here have been processed for an inline cavity and therefore the compatibility with the construction architecture is verified.
As far as the effect of detuning on the cavity is concerned, it has been verified that the bistability threshold also increases with $\Delta\omega$, since an higher resonance shift is expected to visualize the transition.

4.2.3 Limits of convergence between static and dynamical model

The following will discuss to what extent the static and dynamical models converge. In order to do so, two cases will be analysed:

1. **Case 1**: Ramp input signal within [0.01:0.5] mW for both models.
2. **Case 2**: Ramp input signal within [0.01:0.5] mW for the static model and a sinusoidal signal with amplitude of 0.5 mW and variable period for the dynamical model.

Case 1

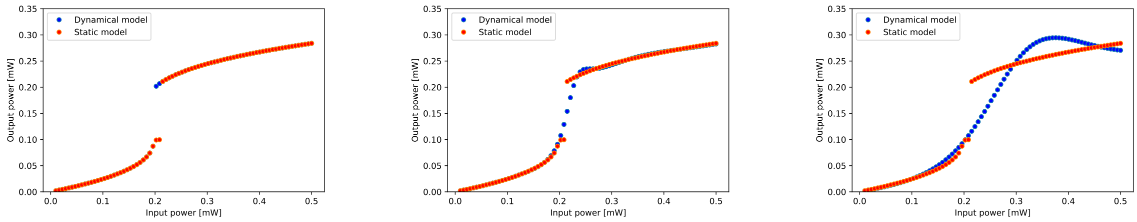


(a) $f = 100MHz$, $\Delta\omega = -0.1^{12}rad/s$ **(b)** $f = 5GHz$, $\Delta\omega = -0.1^{12}rad/s$ **(c)** $f = 10GHz$, $\Delta\omega = -0.1^{12}rad/s$

Figure 4.3: Static model validity range in time domain for fixed input signals with different period of time.

As can be seen in figure 4.3, time domain simulations has been implemented proposing different values for the period of input signal in the cavity. In fact this parameters does not influence the characteristic of the static model, being this time independent and characterized by a single solution for each input value. On the contrary, the dynamical

model is strongly influenced by the period of the input because a dynamic resolution is set for the motion equation that describes the evolution of the mode amplitude in the cavity, eq. 3.8.
The purpose of this comparison is to characterize the error that is obtained by choosing a frequency for the input signal too high, in fact this would involve working outside the static condition of immediate rearrangement of the carriers. The usefulness of this information at a practical level is that through this comparison it will be possible to know how many signals per second it will be possible to send into the cavity in order to obtain a stable and immediate transition.
From figure 4.3 it can be understood that the PhC presents a practically instantaneous response in the MHz range, figure 4.3a. Therefore the model has been simulated up to the GHz domain and it has been observed that the response of the cavity converges with the static solutions up to around 5 GHz[2]. In fact, continuing to increase the frequency of the signal (figure 4.3c) the dynamical model substantially diverges from the static limit.
In order to further verify these assumptions the static solutions for the same power range have been analyzed as visible in figure 4.4 and the considerations obtained previously have been verified.



(a) $f = 100MHz$, $\Delta\omega = -0.1^{12} rad/s$

(b) $f = 5GHz$, $\Delta\omega = -0.1^{12} rad/s$

(c) $f = 10GHz$, $\Delta\omega = -0.1^{12} rad/s$

Figure 4.4: Static model validity range for fixed input signals with different period of time.

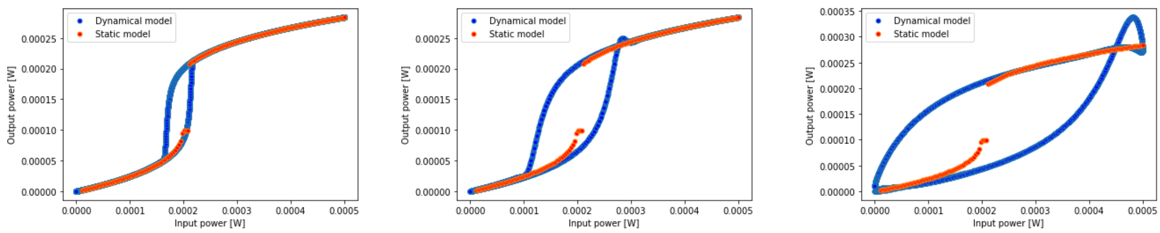
Starting from the 4.4a image it is possible to see that the convergence between the results of the two models is verified for each solution, as expected from the time domain simulation. This convergence consideration is verified for the whole subsequent frequency range up to 5GHz, fig 4.4b where it is possible to notice a slight discrepancy between the two models solutions. Continuing to increase the frequency of the signal the solutions begin to diverge because the cavity has not yet reached a stable solution, as visible in 4.4c. There is therefore convergence between time domain analysis and static solutions.

Case 2

It is important to underline that the use of a sinusoidal input for the dynamic model necessarily involves the formation of a hysteresis cycle, for this reason ascending and descending ramp input has been used for the static model. However, as already pointed

[2]In particular it has been observed that a good level of convergence is achieved up to 5.7 GHz.

out previously, the static model for the moment is not able to visualize this characteristic, so only a partial comparison will be possible.



(a) $f = 10MHz$, $\Delta\omega = -0.1^{12} rad/s$

(b) $f = 100MHz$, $\Delta\omega = -0.1^{12} rad/s$

(c) $f = 1GHz$, $\Delta\omega = -0.1^{12} rad/s$

Figure 4.5: Static model validity range for a sinusoidal input signals with different period of time.

As can be noted, since the photonic crystal cavity is a memory device, the response will depend on the history of the input signal, and therefore on its shape. Particularly for a sufficiently slow signal, i.e. in the order of MHz, the response of the cavity to the signal is almost instantaneous. This behaviour can therefore be assumed as a static response by the cavity. This characteristic is clear when looking at figure 4.5a where the solutions for the two models are perfectly convergent.
By means of the previous consideration it is immediate to ask therefore within what limits the static model is still verified, and observing the figure 4.5c can be noticed that the solutions tend to maintain a convergence up to about 1 GHz, a value from which the response of the cavity is too fast and the static approximation is no longer verified.

Chapter 5

Conclusion and Perspectives

In this work a static model capable of describing a photonic crystal cavity resonator has been developed and analysed and its limits of validity have been studied. Through the considerations obtained here it is intended for the future to correct the gaps in the model realized on photontorch and to try to implement this through laboratory measurements in order to verify the expected results from the simulations. Finally, the goal is also to use this model within a neural network as a decision function and to verify the efficiency of the training. In any case, the approach proposed here of all optical gating applied to the non-linear phenomena of PhC is highly promising both for the high degree of efficiency in terms of energy and sampling and for the intrinsic characteristics related to these devices, and therefore once defined a good level of familiarity with the experimental set up, the possibility to develop this technology within a context of a photonic neural network seems to be a reality far from improbable and the advantages considerable. In fact, of fundamental importance is to remember that such a structure for the characterization of PhC currently exists and the possibility of using this technology in extremely favourable energy conditions is verified.

Appendix A

Here are reported the values experimentally obtained by the group and that have been used for the development of the static model, in particular here GaAs is taken into account:

Parameter	Value	Units
ω	193	2πTHz
Γ	10^{10}	$2\pi s^{-1}$
Γ_C	$1.25 \cdot 10^{11}$	$2\pi s^{-1}$
Γ_A	$4 \cdot 10^{9}$	$2\pi s^{-1}$
V_c	$5.73 \cdot 10^{-20}$	m^3

Table A.1: Model parameters

From article [34] also other materials have been examined, so for more details we recommend to see them.

Bibliography

[1] Floris Laporte, Joni Dambre, and Peter Bienstman. «Highly parallel simulation and optimization of photonic circuits in time and frequency domain based on the deep-learning framework pytorch». In: *Scientific reports* 9.1 (2019), pp. 1–9 (cit. on pp. 2, 19).

[2] Nabil H Farhat, Demetri Psaltis, Aluizio Prata, and Eung Paek. «Optical implementation of the Hopfield model». In: *Applied optics* 24.10 (1985), pp. 1469–1475 (cit. on p. 2).

[3] Demetri Psaltis, David Brady, and Kelvin Wagner. «Adaptive optical networks using photorefractive crystals». In: *Applied Optics* 27.9 (1988), pp. 1752–1759 (cit. on p. 2).

[4] Hsin-Yu Sidney Li, Yong Qiao, and Demetri Psaltis. «Optical network for real-time face recognition». In: *Applied Optics* 32.26 (1993), pp. 5026–5035 (cit. on p. 2).

[5] Ken-ichi Kitayama, Masaya Notomi, Makoto Naruse, Koji Inoue, Satoshi Kawakami, and Atsushi Uchida. «Novel frontier of photonics for data processing—Photonic accelerator». In: *APL Photonics* 4.9 (2019), p. 090901 (cit. on p. 3).

[6] Nicholas C Harris et al. «Linear programmable nanophotonic processors». In: *Optica* 5.12 (2018), pp. 1623–1631 (cit. on p. 3).

[7] J Feldmann, N Youngblood, C David Wright, H Bhaskaran, and WHP Pernice. «All-optical spiking neurosynaptic networks with self-learning capabilities». In: *Nature* 569.7755 (2019), pp. 208–214 (cit. on p. 3).

[8] Paul A Merolla et al. «A million spiking-neuron integrated circuit with a scalable communication network and interface». In: *Science* 345.6197 (2014), pp. 668–673 (cit. on p. 5).

[9] Thomas M Mitchell et al. *Machine learning*. 1997 (cit. on p. 5).

[10] Yung-Yao Chen, Yu-Hsiu Lin, Chia-Ching Kung, Ming-Han Chung, I Yen, et al. «Design and implementation of cloud analytics-assisted smart power meters considering advanced artificial intelligence as edge analytics in demand-side management for smart homes». In: *Sensors* 19.9 (2019), p. 2047 (cit. on p. 5).

[11] Geoffrey W Burr, Pritish Narayanan, Robert M Shelby, Stefano Ambrogio, Hsinyu Tsai, Scott L Lewis, and Kohji Hosokawa. «Neuromorphic devices and architectures for next-generation cognitive computing». In: *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE. 2017, pp. 1–4 (cit. on p. 5).

[12] Kevin P Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012 (cit. on p. 6).

[13] Wolfgang Maass. «Networks of spiking neurons: the third generation of neural network models». In: *Neural networks* 10.9 (1997), pp. 1659–1671 (cit. on p. 6).

[14] Nisheeth Shah. «Left and Right Neural Networks -Inspired by Our Bicameral Brains». In: (June 2019) (cit. on p. 7).

[15] Herbert Jaeger. «The "echo state" approach to analysing and training recurrent neural networks-with an erratum note». In: *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report* 148.34 (2001), p. 13 (cit. on p. 7).

[16] Wolfgang Maass, Thomas Natschläger, and Henry Markram. «Real-time computing without stable states: A new framework for neural computation based on perturbations». In: *Neural computation* 14.11 (2002), pp. 2531–2560 (cit. on p. 7).

[17] Jochen J Steil. «Backpropagation-decorrelation: online recurrent learning with O (N) complexity». In: *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541)*. Vol. 2. IEEE. 2004, pp. 843–848 (cit. on p. 7).

[18] Andrew Katumba, Xin Yin, Joni Dambre, and Peter Bienstman. «A neuromorphic silicon photonics nonlinear equalizer for optical communications with intensity modulation and direct detection». In: *Journal of Lightwave Technology* 37.10 (2019), pp. 2232–2239 (cit. on pp. 8, 10).

[19] Manjari S Kulkarni and Christof Teuscher. «Memristor-based reservoir computing». In: *2012 IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH)*. IEEE. 2012, pp. 226–232 (cit. on p. 8).

[20] Henry O Sillin, Renato Aguilera, Hsien-Hang Shieh, Audrius V Avizienis, Masakazu Aono, Adam Z Stieg, and James K Gimzewski. «A theoretical and experimental study of neuromorphic atomic switch networks for reservoir computing». In: *Nanotechnology* 24.38 (2013), p. 384004 (cit. on p. 8).

[21] Andrew Katumba, Matthias Freiberger, Floris Laporte, Alessio Lugnan, Stijn Sackesyn, Chonghuai Ma, Joni Dambre, and Peter Bienstman. «Neuromorphic computing based on silicon photonics and reservoir computing». In: *IEEE Journal of Selected Topics in Quantum Electronics* 24.6 (2018), pp. 1–10 (cit. on p. 8).

[22] Georg Holzmann and Helmut Hauser. «Echo state networks with filter neurons and a delay&sum readout». In: *Neural Networks* 23.2 (2010), pp. 244–256 (cit. on p. 8).

[23] Russell Stuart, Norvig Peter, et al. *Artificial intelligence: a modern approach.* 2003 (cit. on p. 9).

[24] Geoffrey E Hinton, Terrence Joseph Sejnowski, Tomaso A Poggio, et al. *Unsupervised learning: foundations of neural computation.* MIT press, 1999 (cit. on p. 9).

[25] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. «Reinforcement learning: A survey». In: *Journal of artificial intelligence research* 4 (1996), pp. 237–285 (cit. on p. 9).

[26] Paul A Rutecki. «Neuronal excitability: voltage-dependent currents and synaptic transmission.» In: *Journal of clinical neurophysiology: official publication of the American Electroencephalographic Society* 9.2 (1992), pp. 195–211 (cit. on p. 11).

[27] AL Hodgkin. «Huxley». In: *AF AND KATZ, B.: Ionic currents underlying activity in the giant axon of the squid. Arch. Sri. Physiol* 3 (1949), pp. 129–150 (cit. on p. 11).

[28] Grégory Moille, Sylvain Combrié, and Alfredo De Rossi. «Nanophotonic Approach to Energy-Efficient Ultra-Fast All-Optical Gates». In: *Green Photonics and Electronics*. Springer, 2017, pp. 107–137 (cit. on pp. 12, 17).

[29] Martin Fiers. «Nanophotonic reservoir computing using photonic crystal cavities». PhD thesis. Ghent University, 2013 (cit. on p. 12).

[30] Chad Husko, Alfredo De Rossi, Sylvain Combrié, Quynh Vy Tran, Fabrice Raineri, and Chee Wei Wong. «Ultrafast all-optical modulation in GaAs photonic crystal cavities». In: *Applied Physics Letters* 94.2 (2009), p. 021111 (cit. on p. 13).

[31] Yoshihiro Akahane, Takashi Asano, Bong-Shik Song, and Susumu Noda. «Fine-tuned high-Q photonic-crystal nanocavity». In: *Optics Express* 13.4 (2005), pp. 1202–1214 (cit. on p. 14).

[32] Gregory Moille, Sylvain Combrié, and Alfredo De Rossi. «Modeling of the carrier dynamics in nonlinear semiconductor nanoscale resonators». In: *Physical Review A* 94.2 (2016), p. 023814 (cit. on p. 14).

[33] Hermann A Haus and Weiping Huang. «Coupled-mode theory». In: *Proceedings of the IEEE* 79.10 (1991), pp. 1505–1518 (cit. on p. 14).

[34] Lea Constans, Sylvain Combrié, Xavier Checoury, Grégoire Beaudoin, Isabelle Sagnes, Fabrice Raineri, and Alfredo De Rossi. «III-V/Silicon hybrid nonlinear nanophotonics in the context of on-chip optical signal processing and analog computing». In: *Frontiers in Physics* 7 (2019), p. 133 (cit. on pp. 15, 18, 19, 26).

[35] C Manolatou, MJ Khan, Shanhui Fan, Pierre R Villeneuve, HA Haus, and JD Joannopoulos. «Coupling of modes analysis of resonant channel add-drop filters». In: *IEEE journal of quantum electronics* 35.9 (1999), pp. 1322–1331 (cit. on p. 16).

[36] HM Gibbs, SL McCall, and TNC Venkatesan. «Differential gain and bistability using a sodium-filled Fabry-Perot interferometer». In: *Physical Review Letters* 36.19 (1976), p. 1135 (cit. on p. 17).