Universitat
de les Illes Balears

DOCTORAL THESIS
2023

A time-delay reservoir computing neural network based on a single microring resonator with external optical feedback

Giovanni Donati

Universitat
de les Illes Balears

DOCTORAL THESIS
2023

Doctoral Programme in Physics

A time-delay reservoir computing neural network based on a single microring resonator with external optical feedback

Giovanni Donati

Thesis Supervisor: Apostolos Argyris
Thesis Supervisor: Claudio R. Mirasso
Thesis Supervisor: Lorenzo Pavesi
Thesis tutor: María Rosa López Gonzalo

Doctor by the Universitat de les Illes Balears and the University of Trento

Abstract

Artificial intelligence is a new paradigm of information processing where machines emulate human intelligence and perform tasks that cannot be done with standard computers. Neuromorphic computing is in particular inspired by how the brain computes. Large network of interconnected neurons whose synapses are varied during a learning phase, and where the information flows in parallel throughout different connections. Photonics platforms represent an interesting possibility where to implement neuromorphic processing schemes, exploiting light and its advantages in terms of speed, low energy consumption and inherent parallelism via wavelength division multiplexing. In particular, a candidate playing a diversity of key roles in integrated networks is the microring resonator. In silicon photonics, the microring resonator can implement the strength of a synapse, the spiking emission of a biological neuron, and it can exhibit a fading memory based on its multiple linear and nonlinear dynamical timescales. This manuscript presents an overview of the main applications of silicon microring resonators in neuromorphic silicon photonics, and then focuses on its implementation in a processing scheme, named time delay reservoir computing (RC). Time delay RC is a hardware-friendly approach by which implement a large neural network, where this is folded in the nonlinear dynamical response of only one physical node, such as a dynamical system with delay feedback. The manuscript illustrates, both numerically and experimentally, how to make time delay RC exploiting the linear and nonlinear dynamical response of a silicon microring resonator. The microring is coupled to an external optical feedback and the results on a diversity of time series prediction tasks and delayed-boolean tasks are presented. Numerically, it is shown that the microring nonlinearities can be exploited to improve the performance on prediction tasks, such as the Santa Fe and Mackey Glass ones. Experimentally, it is shown how the network can be set to solve delayed boolean tasks with error-free operation, at 12 MHz operational speed, together with possible upgrades and alternative implementations that can boost its performances.

Resumen

La inteligencia artificial es un nuevo paradigma de procesamiento de información en el que las máquinas emulan la inteligencia humana y realizan tareas que no pueden ser realizadas con ordenadores estándar. La computación neuromórfica está particularmente inspirada en cómo el cerebro realiza cálculos. Consiste en una gran red de neuronas interconectadas cuyas sinapsis varían durante una fase de aprendizaje, y donde la información fluye en paralelo a través de diferentes conexiones. Las plataformas fotónicas representan una interesante posibilidad para implementar esquemas de procesamiento neuromórfico, aprovechando las ventajas de la luz en términos de velocidad, bajo consumo de energía e inherente paralelismo a través de la multiplexación por división de longitud de onda. En particular, un candidato que desempeña una diversidad de roles clave en redes integradas es el micro-anillo resonador. En la fotónica de silicio, el micro-anillo resonador puede implementar la intensidad sináptica, la emisión de pulsos de una neurona biológica, y puede exhibir una memoria que decae con el tiempo basada en sus múltiples escalas temporales dinámicas lineales y no lineales. Esta tesis presenta una visión general de las principales aplicaciones de los resonadores de anillo microscópicos de silicio en la fotónica neuromórfica de silicio y se centra en su implementación en un esquema de procesamiento llamado time delay reservoir computing (RC). Time delay RC es un enfoque favorable para el hardware mediante el cual se implementa una gran red neural, a través de la respuesta dinámica no lineal de solo un nodo físico, como un sistema dinámico sujeto a retroalimentación. Este trabajo ilustra, tanto numérica como experimentalmente, cómo realizar la computación en time delay RC utilizando la respuesta dinámica lineal y no lineal de un resonador de anillo microscópico de silicio. El micro-anillo resonador está acoplado a una retroalimentación óptica externa y se presentan los resultados de una diversidad de tareas de predicción de series temporales y tareas booleanas retrasadas. Numéricamente, se muestra que las no-linealidades del micro-anillo resonador se pueden aprovechar para mejorar el rendimiento en tareas de predicción, como las de Santa Fe y Mackey Glass. Experimentalmente, se muestra cómo la red se puede configurar para resolver tareas booleanas retrasadas sin errores, a una velocidad operativa de 12 MHz, junto con posibles mejoras e implementaciones alternativas que pueden aumentar su rendimiento.

Resum

La intel·ligència artificial és un nou paradigma de processament de la informació en què les màquines emulen la intel·ligència humana i realitzen tasques que no es poden dur a terme amb ordinadors estàndard. La computació neuromòrfica s'inspira particularment en com el cervell calcula. Es tracta d'una gran xarxa de neurones interconnectades, les sinapsis de les quals varien durant una fase d'aprenentatge, i on la informació circula en paral·lel a través de diferents connexions. Les plataformes fotòniques representen una possibilitat interessant per implementar esquemes de processament neuromòrfic, aprofitant la llum i els seus avantatges en termes de velocitat, baix consum d'energia i el paral·lelisme inherent mitjançant multiplexació per divisió de longitud d'ona. En particular, un candidat que juga una diversitat de rols clau en xarxes integrades és el ressonador d'anell microscòpic. En la fotònica de silici, el ressonador d'anell microscòpic pot implementar la intensitat d'una sinapsi, l'emissió d'espigues d'una neurona biològica, i pot exhibir una memòria que decau amb el temps basada en les seves múltiples escales de temps dinàmiques lineals i no lineals. Aquest treballo presenta una visió general de les principals aplicacions dels resonadors d'anell microscòpics de silici en la fotònica neuromòrfica de silici, i després es centra en la seva implementació en un esquema de processament anomenat time delay reservoir computing (RC). Time delay RC és un enfocament amigable per la maquinària mitjançant el qual s'implementa una gran xarxa neural a través de la resposta dinàmica no lineal d'un únic node físic, com un sistema dinàmic amb retroalimentació de retard. Aquest treball il·lustra, tant numèricament com experimentalment, com fer la computació en dipòsit de retard temporal aprofitant la resposta dinàmica lineal i no lineal d'un ressonador d'anell microscòpic de silici. El microanell està acoblat a una retroalimentació òptica externa i es presenten els resultats de diverses tasques de predicció de sèries temporals i tasques booleans retardades. Numèricament, es demostra que les no-linealitats del microanell es poden aprofitar per millorar el rendiment en tasques de predicció, com ara les de Santa Fe i Mackey Glass. Experimentalment, es mostra com la xarxa es pot configurar resoldre tasques booleans retardades sense errors, a una velocitat operativa de 12 MHz, juntament amb possibles actualitzacions i implementacions alternatives que poden augmentar el seu rendiment.

Contents

Chapter 1

Introduction

In the last decades, the field of artificial intelligence (AI) has notably increased, with applications in automotive, face recognition, time series prediction, robotics, and medicine. Differently from conventional computers that are specifically programmed with a precise sequence of instructions to solve a problem, AI algorithms learn how to solve it. Learning is one of the keywords distinguishing these new techniques. Artificial neural networks (ANNs) represent a class of these algorithms that takes inspiration from how a biological brain works. In ANNs models, an ensemble of artificial neurons is connected via artificial synapses to form a network with many diverse topologies. Here, the information can flow in parallel between the neurons (network nodes), and the strength of the connections (weights) changes according to a training algorithm. A learning phase is needed to optimize the network (i.e. change the connections) to solve a particular task, often requiring large training datasets. For this reason, when artificial neural network algorithms are executed in an electronic platform, specialized hardware like Graphics Process Units (GPUs) or tensor process units (TPUs) typically engaged [1]. The latter are able to parallelize the computation, accommodating the intrinsic parallelism of neural network architectures and speeding up their computation.

Neuromorphic photonics represents an alternative and attractive framework where ANNs can be implemented. Its main advantages relate to the fast speed of light and lower power consumption. Furthermore, the intrinsic parallelism demanded by neural networks can be realized here via different spatial photonics nodes, and extends even further to the wavelength, as multiple signals encoded on different wavelength channels can travel in the same

propagating medium. Among others, time-delay reservoir computing (RC) architectures have been widely investigated in photonics. The RC [2][3] concept, essentially treats an ANN as an ensemble of recurrently connected nodes with a *fixed* connection strength and only optimizes a subset of connections, to solve a particular task. The conceptual brake-through of time-delay RC is that the network, rather than being composed of many spatial distinct nodes, can be folded in the nonlinear dynamics of only one physical node [4]. This simplification from a physical implementation point of view, paved the way for several realizations of time delay RC in photonics platforms. Examples of single nodes that are reported in the literature are semiconductor lasers, semiconductor optical amplifiers, feedback loops made by optical fiber, and Mach-Zehnder interferometers, among others.

In my research work, I investigated time-delay RC, realizing a novel hardware based on a silicon microring resonator coupled to an external optical feedback as a nonlinear node. The advantages that microring resonators bring are connected, among others, to their compact footprint, and consequently their integration in photonic platforms. Silicon photonics is chosen leveraging on the progress made by the electronics industry, which allows fabrication in large volumes and operating photonics circuitry in synergy with electronics ones. When the microrings are designed with a high-quality factor, fully passive nonlinearities become also available in silicon. The manuscript is organized in the following chapters.

- **Chapter 2** presents an overview of the different types of ANN models, and the main building blocks of silicon photonics, that allow for their implementation.

- **Chapter 3** presents the modeling of the linear and nonlinear dynamics of a single silicon microring resonator. These will serve to introduce the role that microring resonators have in integrated neural networks. Their linear filtering properties serve as input weights of a photonics neuron, while their nonlinear dynamics serve as a neuron nonlinear activation function.

- **Chapter 4** describes the model of a silicon microring resonator coupled to an external optical feedback and its implementation in a time-delay RC architecture. The system is tested on different benchmark tasks with diverse memory requirements to grasp insights about its computational properties. The findings here are mainly three: first, 4 system

parameters, such as the average input optical power, the optical detuning between the input and the microring resonance wavelengths, the phase and the amplitude of the feedback signal, emerge as critical in determining the performance on these tasks. Second, the microring free carrier nonlinearity is successfully exploited to improve the performances on the benchmark tasks. Third, it is observed that the inertia induced by the microring nonlinearity can also act as a source of memory in the system, where past information injected into the microring is stored in the actual amount of free carriers and temperature in the waveguide.

- **Chapter 5** describes the optical setup for the experimental investigation. A microring resonator in an add-drop filter configuration is specifically designed and coupled to an optical fiber that serves as external optical feedback. This feedback implementation brings experimental unexpected complications, as it introduces phase shifts in the delayed optical signal due to thermal, vibrational, and phonic environmental noise. Most of the experimental effort is devoted to the stabilization of the system against these sources of noise. An optoelectronic controller developed for the stabilization of the system against phase noise and the setup assembled for this experiment are described.

- **Chapter 6** describes the experimental results obtained by testing the microring with an external feedback loop in linear and nonlinear delayed boolean benchmark tasks. The results are compared to other implementations, including the single microring resonator without a feedback loop.

- **Chapter 7** is dedicated to the future perspectives of this work. It presents a fully integrated and upgraded version of the microring with an external feedback loop system, already designed and fabricated. Then, it provides some insights about how to extend the application of the system as a fiber sensor. Finally, it proposes the prototype of a feed-forward neural network for sensing applications.

- **Chapter 8** leads to the final conclusions of the work.

Chapter 2

Theoretical background

2.1 Artificial neural networks

ANNs are inspired by their biological counterparts and aim at providing machines with parallel information capabilities that vaguely mimic those of the brain. An ANN can be represented as an ensemble of units - named nodes or artificial neurons - coupled by connections of different strength (dendrites and synapses). In a simplified model, sketched in Fig. 2.1(a), a biological neuron accumulates the input signals from the dendrites into the soma and, if a threshold potential is overcome, it fires a spike wave potential along the axon, that will then reach subsequent neurons connected by various synapses. Similarly, an artificial neuron has several parallel inputs x_i that are weighted and then summed to emulate the integration phase within the soma, as sketched in 2.1(b). The artificial neuron response, named activation (a), is spread to other neurons and is evaluated as $a = f(\sum_i w_i x_i + b)$, where w_i are the weights setting the relative importance between the input channels, b is a bias term and f is the neuron activation function. The network learns how to solve a task by tuning the interaction between the neurons, i.e. the internal weights and bias values.

The evolution of ANN models can be described in terms of three generations [5]. The first generation is based on the concept of the perceptron model, which was introduced in the 1950s and 1960s [6] [7]. The perceptron refers to a single artificial neuron scheme, as in Fig. 2.1(b), equipped with a

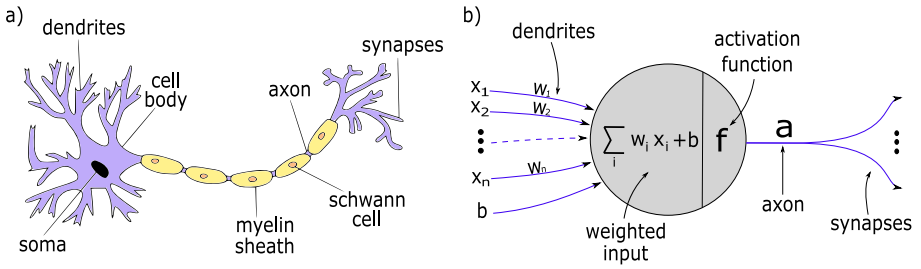

Figure 2.1: Sketch of a) biological neuron and b) artificial neuron, that highlights their common features: integration of the input signals and sharing of the output response with subsequently connected neurons.

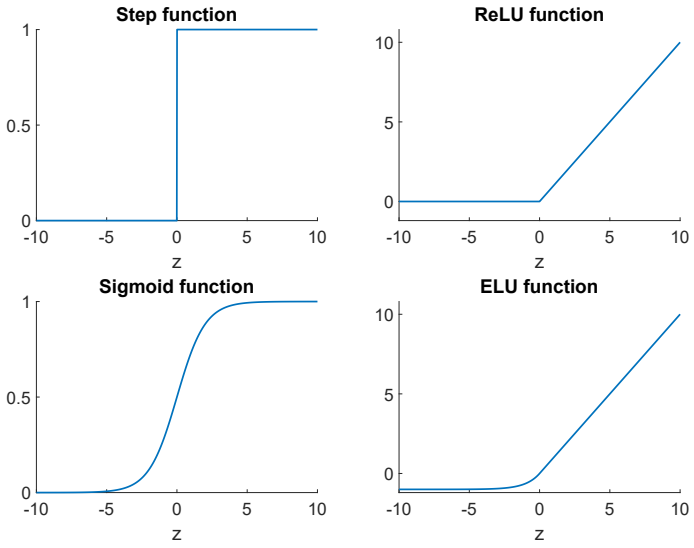

Figure 2.2: Common nonlinear activation functions adopted in ANN models, as a function of the weighted input to the neuron $z = \sum_i w_i x_i + b$.

step-like activation function:

$$f(\vec{w} \cdot \vec{x}) = \begin{cases} 1 & \text{if} \quad \vec{w} \cdot \vec{x} + b > 0; \\ 0 & \text{otherwise;} \end{cases} \tag{2.1}$$

where $\cdot$ indicates the inner product in the Euclidean space and $\vec{x}$, $\vec{w}$ are the input and weight vectors, respectively. The perceptron was initially inspired by the "all-or-none" spiking response of biological neurons. This model can recognize two different categories of inputs by testing whether $f(\vec{w} \cdot \vec{x})$ is

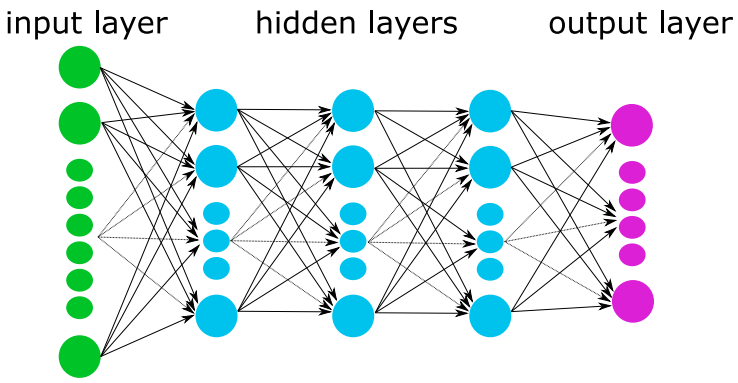

Figure 2.3: Scheme of a multilayer perceptron fully connected to realize a FFNN. While the design of the input and output layer is generally dictated by the problem that the network is designed to solve, the size of the hidden layers can be different.

positive or null, and is the basic unit of a variety of first generation neural networks, like the multi-layer perceptron models.

In the second generation of ANNs, the activation function is extended from the step function to the continuous domain, providing the networks with an internal analog representation of the information. Examples of activation functions are illustrated in Fig. 2.2.

Finally, in the third generation of ANNs, neurons are modeled with a more detailed internal dynamic proper of biological neuron models (section 2.4). The result is a signal-encoding paradigm where information is encoded at the time of binary spike events.

Independently on the generation of neurons adopted, specific topologies emerge when artificial neurons are linked together. ANNs topologies can be essentially classified within two main categories: feedforward and recurrent.

Feedforward neural networks In feed-forward ANNs neurons are grouped into layers, as presented in Fig. 2.3. The information to be processed is represented by the state of neurons belonging to the input layer. This information is then propagated unidirectionally through the network, according to the internal connections, until it reaches the output layer. The state of the output neurons encodes the network response to a certain input. The dimension of the input and output layers is generally dictated by the particular problem at hand. For example, in image classification tasks, the input

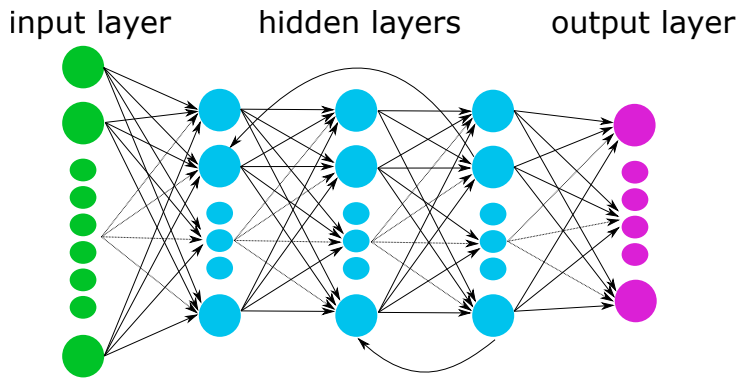

Figure 2.4: A recurrent neural network is characterized by internal recursive connections, as compared to feed-forward ones.

layer can represent the value of the pixels in the image with an equivalent number of neurons. While the dimension of the output layer may represent the total number of output classification classes where the image needs to be categorized, the dimension of each hidden layer and their number is not guaranteed by specific a priori rules and typically requires testing the network on different configurations [8]. Deep FFNNs include at least one hidden layer.

Peculiar of FFNNs is that the information flows unidirectionally from the input to the output layer. There are no loops in the network that allow the information to propagate back, and mix to the propagation of future input signals. As a consequence, FFNNs are applied on tasks that do not require memory, such as pattern recognition and computer vision.

Recurrent neural networks Recurrent Neural Networks (RNNs) include recursive connections within the network, as shown in Fig. 2.4, so that the input information at a given time is preserved within the system for a certain period of time. Due to this, the unique input-output correspondence of FFNNs holds no more, since the network response to a certain input is also affected by the history of the system.

Essentially any function involving recurrence can be considered a recurrent neural network. The classical form of a recurrent dynamical system can be expressed as:

$$\vec{a}_t = f(\vec{a}_{t-1}, \vec{x}_t, \vec{w}), \tag{2.2}$$

where $\vec{a}_t$ are the state of the units at a time t, $\vec{x}_t$ is an element of the input sequence, and $\vec{w}$ are the weights related to the internal connections. Eq. 2.2 is recurrent, since the state $\vec{a}$ at time t refers back to $\vec{a}$ at time $t-1$, and similarly this last. Hence, RNNs find applications in memory demanding tasks, such as handwriting recognition [9], speech recognition [10], and time series prediction [11].

2.2 Training a neural network

Learning is the key property that differentiates machine learning algorithms, such as ANNs, with respect to algorithms specifically developed for a particular problem. In the latter case, a precise list of instructions needs to be provided to the machine in advance, describing its response to the incoming input signal. On the other hand, learning algorithms are designed with internal parameters that adapt from experience. These parameters, so far called weights, are optimized to minimize an output error of the network with respect to a desired output.

Learning strategies can be categorized into three main classes: supervised, unsupervised, and reinforcement learning.

In supervised learning the network is trained with examples (x_i, y_i), with target values y_i a priori known for each input data x_i. If initially an untrained network will output values $o_i \neq y_i$, during training the internal weights are adapted to minimize the error between y_i and o_i. The second generation ANNs, relying on differentiable activation functions, were the first to introduce support to supervised learning algorithms based on gradient descent [12]. A famous example is given by the back-propagation algorithm, previously proposed in 1960 [13], and then adapted for the training of deep ANNs [14]. Back-propagation combined with stochastic gradient descent algorithms rely on a differentiable cost function that is defined to minimize the output layer response and the desired target. The weights and bias of the network are updated according to the derivatives of the cost function with respect to them (with opposite sign), in a chain rule that starts from the last layer and back-propagates towards the first hidden layer. Back-propagation was also adapted for training of RNNs [15].

On the other hand, in unsupervised learning, the targets are not known a priori, and the network is trained to find patterns underlying unlabeled data. Unsupervised learning is particularly useful for at least two reasons: cluster-

ing generation from large amounts of unstructured raw data, and generation of labels useful then for supervised learning.

Reinforcement learning is based on an agent, that must learn behavior through trial-and-error interactions with a dynamically changing environment. At each time step, the agent receives an input signal, considers the actual state of the environment, and takes actions that in turn modify the environment. Reinforcement learning algorithms reward the agent for taking actions that lead to successful states. Reinforcement learning finds applications in game playing [16], robotics, and control [17].

The choice of the input and target datasets in supervised learning, or the strategies of awards delivered in reinforcement learning, finally define the task that the neural network is trained to solve. Once finalized the learning, the network undergoes a test phase, where it is presented with new datasets not seen before. In this phase, the trained weights of the network are no more changed and the task error directly measures how much the features previously learned generalize to the new dataset. Overfitting the training dataset is a possible risk during the training phase, which induces the network to learn insights that are particular to the training dataset but that are not generalizable to new ones.

For large networks, which involve recurrences particularly, the learning phase is time and resource challenging. A new paradigm for the training of RNNs, namely reservoir computing, is described in the next section.

2.3 Reservoir computing

Reservoir computing (RC) is a new paradigm to recurrent ANNs, introduced to simplify their training [2][3]. As the general scheme in Fig. 2.5 illustrates, the recurrent network is treated as a black box whose connection's strength and topology are left unchanged over time. For this reason, the recurrent network is also named reservoir and does not participate in the training process. The reservoir's dynamic is driven by the information to process, which is encoded at the input layer. A linear output layer is then estimated via a weighted sum of the reservoir node states.

Mathematically, the state $\mathbf{a}$ of a reservoir composed of N nodes, can be expressed as:

$$\mathbf{a}(i) = \mathbf{f}(W_{in}\mathbf{x}(i) + W\mathbf{a}(i-1) + \mathbf{b}) \tag{2.3}$$

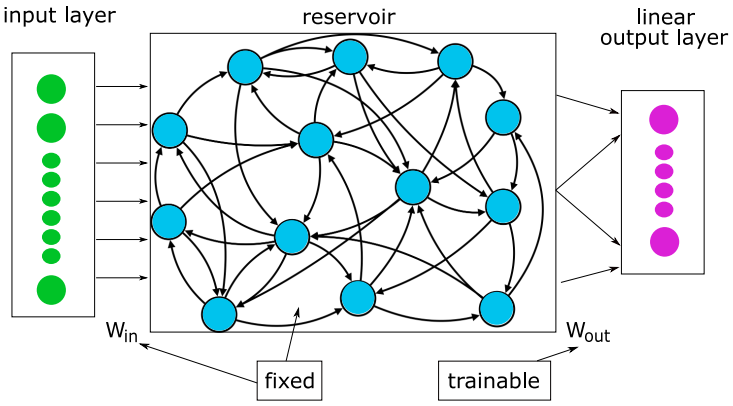

Figure 2.5: Schematic of reservoir computing, consisting of an input layer, a reservoir of interacting nodes with fixed topology and strength, and a final output layer. The training only involves the weights w_{out} of the linear readout layer.

$$\mathbf{y}(i) = W_{out}\mathbf{a}(i) \tag{2.4}$$

where $\mathbf{x}(i) \in \mathbb{R}^{k\times 1}$ is a k-dim input vector, $W_{in} \in \mathbb{R}^{N\times k}$ refers to the input weights, $W \in \mathbb{R}^{N\times N}$ refers to the internal weights of the reservoir, and $f \in \mathbb{R}^{N\times 1}$ describes the nonlinear activation function of each neuron. Then $W_{out} \in \mathbb{R}^{k'\times N}$ are the output weights that project the reservoir neurons to the output layer, whose state identifies the response $\mathbf{y} \in \mathbb{R}^{k'\times 1}$ of the network.

In RC only the output weights W_{out} are trainable parameters. The reservoir serves to nonlinearly project the input information onto a higher dimensional space, given by the number of internal nodes. From the reservoir state, the correct output can be estimated by a simple linear readout layer. As an example, let's consider the boolean XOR task also represented in Fig. 2.6. The XOR problem, in the two-dimensional space created by the possible combinations of input $[(0,0),(1,0),(0,1),(1,1)]$ is not linearly separable, i.e. there is no line able to successfully separate the correspondent targets $[0,1,1,0]$ (encoded as orange and blue colors in Fig. 2.6). Instead, if the input is nonlinearly mapped into a higher dimensional space, for example with an extra dimension (analogous to the reservoir), it becomes possible to find a hyperplane (linear readout, with coefficients given by W_{out}), leading to the correct classification of the 2 classes.

As a consequence of the "black box" approach, different complex physical systems have been proposed to serve as reservoir implementations. An ex-

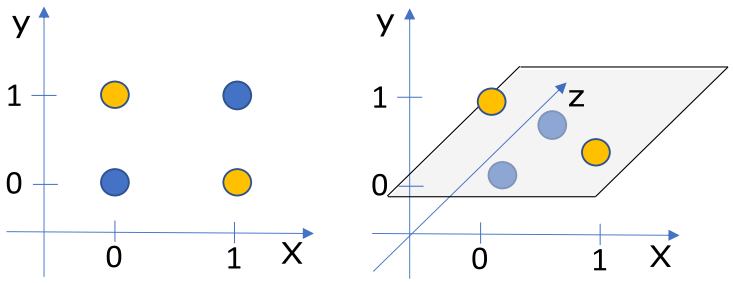

Figure 2.6: XOR problem: in two dimensions, no line can successfully separate the two classes. On the other hand, under an appropriate nonlinear mapping into a higher dimensional space, the problem becomes linearly separable.

ample highlighting the potential of a physical RC system is reported in [18]. In this work, a tank of water is perturbed on the surface, according to the input to process, by 8 LEGO engines. The tank of water is illuminated from the bottom. Light propagates up to the water's surface, where it is scattered according to the wave pattern present at that moment. The scattered light finally reaches a camera positioned above the setup, where is detected. The pixel values are considered the reservoir node states, which are then, offline, linearly combined to generate the output of the network. In the experiment, the reservoir nodes (pixels) are coupled with each other thanks to the wave pattern on the water's surface. This experiment is very instructive, as it shows how physical reservoirs can be beneficial in terms of computational cost and reservoir dimensionality, rather than their simulation: the coupling of the reservoir nodes occurs naturally in the system, exploiting the interaction between water particles, while it would require a large computational power and knowledge if simulated.
Computing systems with multiple reservoirs have also been proposed. For example, in [19], the authors propose a deep architecture that exploits the RC framework at each layer. Different reservoirs across the layers, provide a diversification of the temporal input representation across the layers, thus providing here the full network with different processing timescales.

Time-delay RC Time delay RC is a further simplification of the RC approach, introduced in 2011 by Appeltant at al. [4]. In this approach, the conceptual breakthrough is the possibility to emulate networks with a large

number of neurons, exploiting the nonlinear dynamics of only a single node, such as a dynamical node subject to delay. Within the feedback loop, virtual nodes are defined [4]. A network of virtual nodes is then extracted by sampling the nonlinear response of the single node to a certain input information, and then using the recording of these temporal samples to feed the output layer, as in traditional RC.
The approach is suitable for physical implementations as it minimizes the hardware requirements. Many physical implementations have been investigated in this context, including the one presented in this thesis. A detailed description of this processing scheme will be presented in Chapter 4.

2.4 Spiking neural networks

ANNs, as presented until now, consider the input information from different channels always reaching a neuron synchronously, and then summed. These times are typically discretized by the simulation time step. In Spiking Neural Networks (SNN)s, this restriction is left, so that input signals can reach a neuron delayed with respect to each other, and the synchronicity is no more guaranteed. The neuron is here described as a spiking neuron. Similar to first generation neurons, spiking neurons act as integrate-and-fire units and have binary response. The spiking neuron, however, has an inherent dynamic nature characterized by an internal state which changes with time. This allows ANNs, to mimic more closely biological neural networks, where information can be encoded in the time at which individual spikes are generated, as well as in the firing rate.

The spiking neuron needs to satisfy some specific characteristics:

- Excitability threshold: each neuron fires an action potential (spike) at the time instance its internal state exceeds the neuron threshold.

- Leaky integrating dynamics: the inherently dynamic nature of each neuron allows the integration of sufficiently closed-spaced sub-threshold pulses, eventually leading to its excitation and spike emission.

- Refractory period: after each spike emission, a time interval follows where the neuron relaxes to its rest state. During this time, the neuron is inactive and does not fire, independently of the input it is receiving.

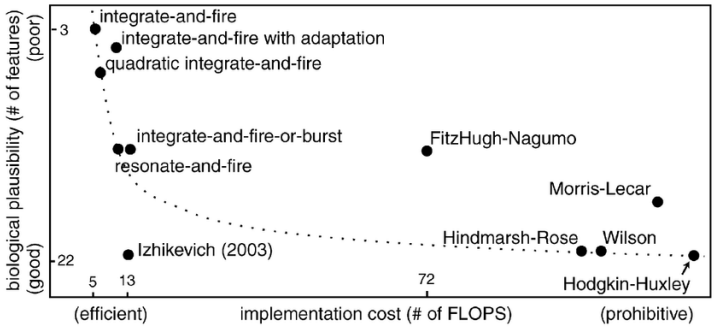

Figure 2.7: Trade-off between biological plausibility and implementation cost in FLOPS per 1ms of simulation time, for several mathematical and biophysical models. Taken from [20].

- Cascadability: the output spike of a neuron needs to be strong enough for subsequent neurons to be excited as well.
- Inhibitory spiking behavior: the arrival of an input stimulus can reduce the probability of spiking.

Spiking neurons can be modeled in different ways, with a number of detailed mathematical and biophisical models developed to quantitatively characterize biological neuron behaviors (Fig. 2.7). The presence of an internal dynamic, when combined with the need to use large networks, leads to a high computational cost of SNNs. This is why first and second generation ANNs, based on simplified neuron models, were first investigated. Nowadays, the computational power available also allows the simulation of large SNNs. The field, referred to as computational neuroscience, studies electrophysiological processes, pattern generation, and the dynamic behavior of groups of neurons. The reader can refer to [21] for a detailed review of the topic.

2.5 Neuromorphic computing

While conventional computer allows the simulation of ANNs models, their centralized processing architecture provides a bottleneck regarding the speed of operation. Central Processing Units (CPUs) operate indeed in a serial way, with continuous access to the cache memory that limits the intrinsic parallelism underlying ANN models. For this reason, electronic hardware

with parallel processing capabilities has also been developed to accelerate the simulation of ANNs. An example is given by Graphic Processing Units (GPUs) that, originally were introduced to boost the performance of video games, were then also exploited for the simulation of ANNs. Another example is given by Tensor Process Units (TPUs), application-specific integrated circuits originally developed by Google in 2017 to improve the costs and performance of ANNs inference in their data centers [1].

Neuromorphic electronic hardware has been also developed. Notable examples in this regard are the SpiNNaker [22] and the IBM's TrueNorth [23], whose computing platforms are based on spiking neural networks and rely on parallel and custom cores with programmable connectivity, for high-efficiency computation. However, electronics implementations face fundamental limits as Moore's law is slowing down [24][25]. Moreover, moving data electronically on metal wires has fundamental bandwidth and energy efficiency limitations.

Photonics represents an attractive and alternative framework for producing accelerated processors where artificial network schemes are combined with the speed of light [26]. Research in neuromorphic photonics encompasses a variety of hardware implementations and neural network types, with both free space, fiber optics, and integrated techniques.

An example of free space implementation is the diffractive deep network proposed in [27]. Here, an input-encoded light propagates through subsequent diffractive layers (with an area of $8cm \times 8cm$), until is collected by ten detectors placed in different positions at the output layer. Each layer, can be seen as an ensemble of secondary wave sources, which integrate the diffracted light from the previous layer, apply a complex modulation, and transmit it to the 'neurons' of the next layer. Differently from a FFNN model, the nonlinearity here only applies at the output layer, via the photodetection square low, and not across the layers, thus highlighting the flexibility of neuromorphic engineering, rather than being limited to strictly emulate ANNs models.

Other examples of free space networks exploit the RC paradigm [28][29]. Spatial light modulators encode here the input information in the optical domain, and, interestingly, the reservoir is just the simple linear optical propagation of light before being nonlinearly detected.

Free space optics for RC is also applied in [30], to address the optical input information towards an integrated matrix of 5×5 Vertical Cavity Surface Emitting Lasers (VCSELs), which serve as a nonlinear reservoir.

On the other hand, the rapid development of Photonic Integrated Cir-

cuits (PICs), provides a promising framework for ANNs. Today there exists a wide diversity of technology platforms for PICs. These include different material systems such as III-V semiconductors, Lithium Niobate, high-index glasses and nitrides, polymers, and silicon. Silicon photonics has the unique advantage to be CMOS compatible and benefits of advanced fabrication technologies proper of the electronics industry. High volume manufacturing at potential low cost is possible, with, at the same time, the possibility to integrate photonic ANNs with electronic functions. This synergy allows for electronic feedback controls and tuning of the photonics components that suit the programmability of the internal weights required by ANNs. Moreover, the high refractive index contrast between the silicon core and the silica cladding used to confine light allows for sub-micrometer waveguide dimensions (waveguide cross-section of $250nm \times 500nm$), lower bends ($< 5\mu m$), and thus for dense functional structures on the surface of the chip, despite real scalability of the networks is still an issue due to propagation losses.

Leveraging these properties, silicon photonics is already the technology adopted in large data centers, to route the increasing global data traffic [31]. It also finds applications in biosensing and diagnostics [32] [33], spectroscopy [34], sensor functions, such as Lidar [35] and appears now as a promising framework for realizing integrated ANNs.

The next section will introduce and briefly review the main building blocks of a silicon photonic network.

2.6 Silicon photonics building blocks

An integrated photonic circuit relies on a multilayer structure as sketched in Fig. 2.8, where a silicon layer is surrounded by a cladding and a substrate, typically made of silicon dioxide (SiO_2). This ensures the possibility to trap light inside the silicon layer, where the optical integrated components are built, and thus guide it through the different structures.

2.6.1 Dielectric waveguides

Dielectric waveguides represent the physical pathways along which light can be driven on the surface of a photonic chip, interconnecting its different optical components. They consist of a silicon core ($n_{si} \approx 3.4$, at $\lambda = 1.55\mu m$),

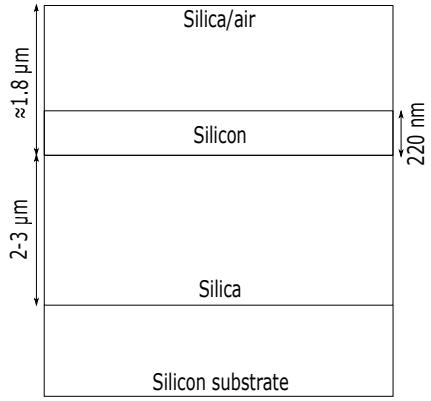

Figure 2.8: Illustration of an SOI wafer cross-section with typical dimensions, having a silicon core and silica cladding and substrate. Also air, for example, can serve as a cladding. The bottom silicon layer is used as substrate of the structure while the optical components are built in the top silicon layer.

surrounded by a cladding material with a lower refractive index, like silica ($n_{SiO_2} \approx 1.55$, at $\lambda = 1.55\mu m$) or air ($n_{air} = 1$). Light is prevented to escape the core thanks to total internal reflection (TIR)

Relying on the TIR principle, dielectric waveguides can be classified according to their shape, as sketched in Fig. 2.9. In the simplest step-index (slab) waveguide light is confined in only one direction at the upper and lower interfaces. Adding lateral boundaries to the core material allows TIR to take place also at the lateral interfaces, thus providing an additional confinement dimension of the optical field. These 2D waveguides are typically named strip and can be both embedded in the substrate or on top of it. In this last case, the core layer can be either etched fully to the bottom oxide layer or partially etched, providing in this last case a rib waveguide. Note that optical fibers, which are widely used in nowadays off-chip long-distance communications, are also a well-known example of channel waveguides with geometrical shape.

Light propagates inside a 2D dielectric waveguide in the form:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}) e^{i\phi(\mathbf{r})} e^{-iwt}, \tag{2.5}$$

with $\mathbf{E}(\mathbf{r})$ and $\phi(\mathbf{r})$ being the complex amplitude and phase, solutions of the Helmholtz equation [36]. Only approximate solution methods, such as finite

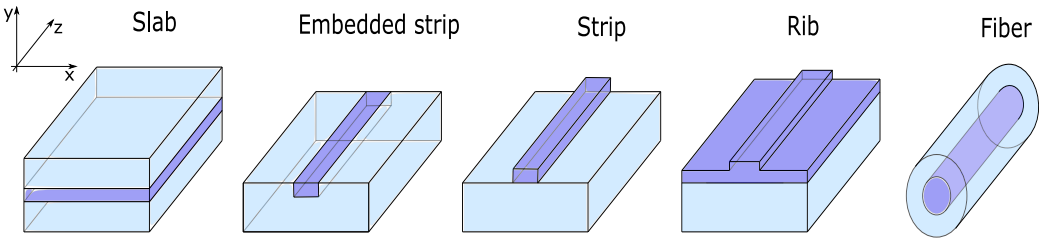

Figure 2.9: Waveguide geometries, with the darker blue region representing the core of the waveguide where light propagates.

element methods (FEM), can be used for an estimation of $\mathbf{E}(\mathbf{r})$ and $\phi(\mathbf{r})$ in 2D-waveguides, while analytical solutions exist for the 1D slab waveguide. In both cases, the finite dimensionality along one or more directions and the necessity to satisfy the continuity of the tangential component of the electric and magnetic fields and their derivative across the interfaces leads to the concept of mode. Assuming a monochromatic wave and that the refractive index does not depend on the propagation direction z (translational invariance), the propagating mode can be described by:

$$\mathbf{E}_{m,n}(x,y,z) = \mathbf{E}_{m,n}(x,y)e^{j\beta_{m,n}z} \tag{2.6}$$

where $E_{m,n}(x,y)$ describes the transverse shape of the field with respect to the propagation direction, $\beta_{m,n}$ is the propagation constant, and m and n are integer numbers arising from the quantization of the wave vector components k_x and k_y. These last indicate the number of nodes that the node exhibit in the correspondent directions.

The propagating field is classified as transverse electric (TE) or transverse magnetic (TM), depending on which field component is transverse to the propagation direction. Examples of TE and TM polarized modes are reported in Fig. 2.10, for a multimode waveguide of dimension $1.6\mu m \times 0.25\mu m$.

Fig. 2.10 also highlights another important characteristic of a polarized mode: the modal confinement factor. This last can be defined as:

$$\Gamma_{c(m,n)} = \frac{\int_{wg} n^2 |E_{m,n}(x,y)|^2 dxdy}{\int n^2 |E_{m,n}(x,y)|^2 dxdy}, \tag{2.7}$$

where at the numerator the integral extends only to the waveguide (wg) area. As the mode order increases, a larger portion of the light profile extends beyond the waveguide core, thus leading to a lower confinement factor. At

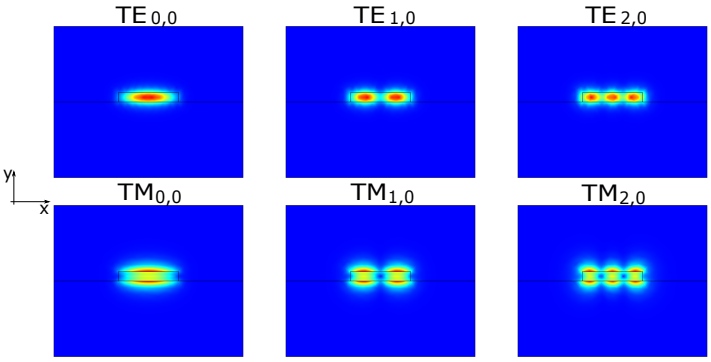

Figure 2.10: Field power distribution for the three lowest order TE and TM polarised modes, in a silicon waveguide of dimension $1.6\mu m \times 0.25\mu m$, surrounded by silica. Simulations performed in Comsol.

the same time, the effective modal index decreases, as the optical field senses less the core refractive index and more the surrounding refractive index.

It is thus convenient to express the propagation constant in terms of the effective refractive index: $\beta_{m,n} = \frac{w}{c} n_{eff,m,n}$, where $n_{eff,m.n}$ is still a complex quantity whose real part is related to the propagation and whose complex part is related to losses. Losses caused the field intensity to exponentially decreases along z as $e^{-\alpha z}$, with α the attenuation coefficient given by $\alpha = \frac{2\pi}{c} Im(n_{eff,m,n})$.

The expressions of the field distribution reported until now rely on the assumption that the field is monochromatic. Otherwise, the dispersion of the material with respect to the wavelengths has to be considered. Assuming a small range of frequencies, the propagation constant can be expanded in series:

$$\begin{aligned} \beta(\omega') &\simeq \beta(\omega) + \frac{\partial \beta(\omega)}{\partial \omega}\Delta\omega + \frac{1}{2}\frac{\partial^2 \beta(\omega)}{\partial \omega^2}\Delta\omega^2 \\ &\simeq \frac{1}{v_{ph}}\omega + \frac{1}{v_g}\Delta\omega + \frac{1}{2}GVD\Delta\omega^2 \end{aligned} \tag{2.8}$$

where v_{ph} is the phase velocity, v_g the group velocity and GVD the group velocity dispersion, defined as following:

$$v_{ph} := \left(\frac{\beta}{\omega}\right)^{-1} \tag{2.9}$$

$$v_g := \left(\frac{\partial \beta}{\partial \omega}\right)^{-1} \tag{2.10}$$

$$GVD := \frac{\partial^2 \beta}{\partial \omega^2} = \frac{\partial}{\partial \omega}\frac{1}{v_g} \tag{2.11}$$

They describe respectively the speed at which the wavefront of each frequency propagates in space, the speed at which the modulation of the wave propagates, and finally how the medium affects the single frequency components of the pulse.
In analogy with the refractive index, also an effective group index is defined as:

$$\begin{aligned} n_{eff}^{g} := \frac{c_0}{v_g} &= \frac{\partial(\omega n_{eff}(\omega)}{\partial \omega}) = n_{eff}(\omega) + \omega \frac{\partial n_{eff}(\omega)}{\partial \omega} \\ &\simeq n_{eff}(\lambda) - \lambda \frac{\partial n_{eff}(\lambda)}{\partial \lambda} \end{aligned} \tag{2.12}$$

where the group velocity is used at the denominator instead of the phase velocity.

Losses in a waveguide are due to both scattering and bulk-absorption processes. Scattering losses are in particular caused by the waveguide sidewalls [37]. Indeed, when the optical mode distribution overlaps with the side walls, the irregularities present here become a source of backscattering, thus attenuating the propagating beam. Sidewalls roughness are the result of fabrication process such as the optical lithography and dry-etching process and can be reduced either by processes aiming to smooth the sidewalls [38] [39], or by optimizing the geometry such to reduce modal field intensity at the side walls [40]. As a result, a lower confinement factor is also an index of higher losses, since more light feels the surface effects. In the example of Fig. 2.10, the waveguide is designed to support TE modes and would present higher propagation losses for TM modes. This is why, experimentally, it is important to properly polarize the optical field before coupling it to the chip.

2.6.2 Couplers

Optical coupling refers to all the operations in which light is injected into a waveguide. Light can be squeezed into the photonic waveguide without

changing its propagation direction, thus using an in-plane approach, or otherwise using a less intuitive out-of-plane approach.

Edge coupling In edge coupling or butt coupling, light is coupled from the waveguide edges. In this case, efficient coupling requires an input field profile overlapping the waveguide core. For example, if the coupling involves a waveguide (core $\sim 1\mu m^2$) and a standard single mode fiber (SMF) 28 (core $\sim 10\mu m^2$), the geometrical mismatch between the modes leads to high insertion losses. In addition, due to the refractive index difference ($n_{S_i} = 3.45$, $n_{S_iO_2} = 1.55$), large Fresnel reflection occurs. In this case, the facets of both the fiber and the waveguide can be engineered to reduce the mismatch. Inverse tapering techniques gradually decrease the width of the waveguide along the direction of light propagation. As the waveguide dimension decreases, the mode becomes less confined, its effective cross-section increases and the effective index reduces to become as similar as possible to the effective index of the fiber mode. Waveguide lengths longer than 100 μm, are required for adiabatic conversion of the waveguide width to few tens of nanometers, while the height remains dictated by the thickness of the silicon layer.

For reducing farther the losses, lensed optical fibers can be used, where one of the two facets is polished (or laser-ablated) into a hemispherical tip, which allows focusing the fiber mode down to a $2\mu m$ diameter spot, for an improved fiber-waveguide mode overlap.

Grating coupling In an alternative way, light is coupled into the waveguide from its top surface through grating couplers. In this case, a grating is written across the waveguide by periodic etching of the waveguide core's surface, thus creating alternating regions of different refractive indices. Light can be diffracted into the waveguide if the grating equation is satisfied. In fact, light can be coupled into the waveguide (or scattered from the waveguide), with an angle θ (with respect to the normal to the grating), provided that the following phase matching condition is satisfied:

$$sin\theta = \frac{\Lambda n_{eff} - \lambda_0}{\Lambda n}, \tag{2.13}$$

where Λ is the etching period, n_{eff} the effective index of the waveguide mode, λ_0 the wavelength of light, and n the refractive index of the medium where

the wave is radiated.

While edge coupling strategies offer better performance, with higher coupling efficiencies and almost flat bandwidth, grating coupling techniques result more practical for accessing and testing several structures on a photonic chip, but with the disadvantage of a limited working band. A detailed review of the topic with the relative efficiencies can be found in [41].

2.6.3 Splitters and combiners

In order to achieve a high density and large connectivity on-chip, strategies to exchange light between waveguides or in-plane waveguide-crossings are desirable. Several components serve this purpose, from directional couplers that simply couple light via evanescent field, to engineered structures such as Y-junctions and Multi-Mode Interference devices.

Directional coupler A first design consists of two waveguides sufficiently close that their evanescent fields overlap. Depending on the separation distance, the coupling length, and the effective index mismatch, the optical field can be exchanged in different amounts between the two waveguides. For example, Fig. 2.11 shows the coupling between two straight waveguides. The waveguides have a silicon core ($220nm \times 480nm$) and are surrounded by silica (substrate) and air (cladding). Light feeds only one waveguide, and is collected at the output port of the other waveguide, as a function of their gap distance, while using a pumping wavelength $\lambda_p = 1550nm$. The result shows that the optical coupling decreases exponentially while increasing the gap width between the two waveguides, in agreement with the exponential decay of the evanescent field out of the waveguide core.

Crossings Other than coupling light between waveguides, there may be the need to cross multiple waveguides minimizing their cross-talk. In a simple direct crossing configuration (Fig. 2.12(a)), often presented as a four-ports, symmetric circuit component, the lateral confinement is lost in the crossing region, causing diffraction of light, cross-talk between the waveguides, and hence optical losses.

Optimized design aim at avoid diffraction in the crossing region by customizing the crossing geometry to suppress the wide-angle component of the

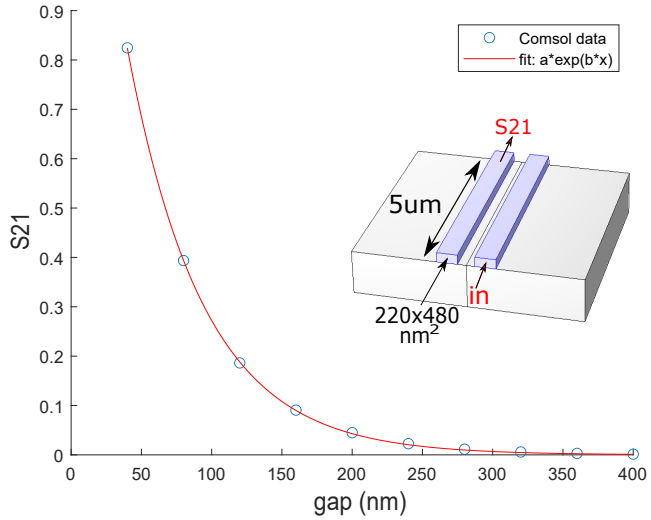

Figure 2.11: Coupling efficiency between two straight silicon waveguides. Light is input to one waveguide and the normalized optical transmission $S21$ is measured at the output of the second waveguide. Simulations performed with Comsol.

mode (Fig. 2.12(b)). In these cases, the mode is evolved near the crossing to a plane wave, for example using an elliptical or parabolic mode expander, combined with an adiabatic tapering that smoothly expands the guided modes, without exciting high order modes.

Another efficient strategy is to etch holes just before the crossing region, obtaining a lens-like structure. In this way, a crossing device has been demonstrated with insertion loss lower than 0.175 dB and crosstalk lower than -37 dB, while maintaining an extremely compact footprint of $\sim 1 \times 1 \mu m^2$ [43].

It is also possible to avoid crossing between waveguides by creating a bridge in the crossing region. Here, light is coupled from the silicon channel into a top layer before the crossing, and then coupled back to the silicon channel after the crossing is surpassed. This is an example of vertical directional coupler method (also named bridge method). The coupling efficiency strongly depends on the top material employed and can be maximized with optical impedance matching. An amorphous silicon layer has been proposed in [44], because of its high refractive index (3.3-3.7) and the possibility to be integrated during SOI fabrication processes. In this work, the authors

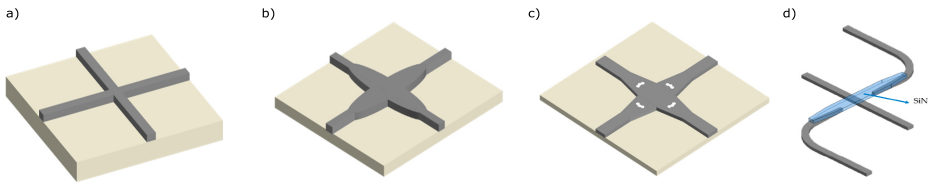

Figure 2.12: Scheme of crossing structures on silicon-on-insulator. a) Direct waveguide crossing. b) Shaped taper waveguide crossing. c) crossing with the holey sub-wavelength grating method. d) crossing with a vertical directional coupler. Taken from [42].

emphasized how tapering both the bottom and top waveguides, within the vertical coupling region, can be important to have larger tolerances in the variations of the amorphous silicon refractive index and satisfy the optical impedance matching condition. Insertion losses as low as 0.2 dB can be obtained in this way. Other materials have been also investigated like polymers [45] and silicon nitride [46].

Y-junction Another need is an optical component that splits or combines the signal on the chip. Y-junctions are an example of waveguide-based structures that, thanks to a specific tapering, allow a single input channel to be equally separated into two output channels, and vice-versa. Optimization algorithms are used to optimize the tapering width. For example, the taper can be first digitalized in a certain number of segments in the direction of light propagation, each one related to a certain width. Then, a figure of merit sets the goal of the structure and, accordingly, the width values of each segment are optimized. The tapered geometry is finally given by the interpolation of these points, as shown in Fig. 2.13 [47].

Crossing designs, Y-junctions, and more in general multimode interferometers are typically designed by starting with an analytically designed structure and then hand-tuning a few parameters. Inverse design methods are algorithms that automatize and extend the hand-tuning process, to a much wider space of possible structures.

Among the many approaches, described in [48], inverse design algorithms share common features. The physical structure is firstly divided into pixel domains, which account for the precision allowed by the fabrication process in depositing the material in a certain position. Then, a real function $\phi(x, y) : \mathbb{R}^2 \to \mathbb{R}$ is defined in the pixel domain, and constrained to $\phi = 0$, at the

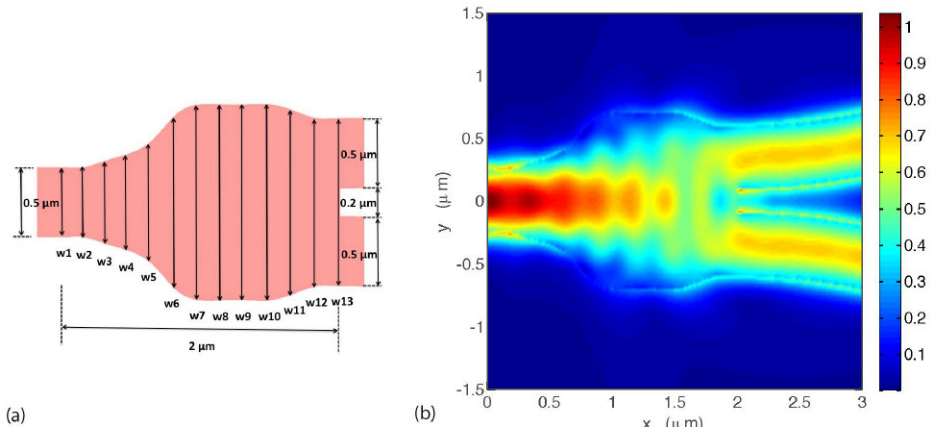

Figure 2.13: a) Sketch of Y-junction layout with tapering optimized in width for each horizontal segment, and b) contour plot of the simulated E-field distribution at 1550 nm wavelength. Taken from [47].

structure's boundary. A permittivity function ϵ is also defined to consider smooth candidate structures consisting of two materials:

$$\epsilon(x,y) = \begin{cases} \epsilon_1 & \text{if} \quad \phi(x,y) \leq 0; \\ \epsilon_2 & \text{if} \quad \phi(x,y) > 0; \end{cases} \tag{2.14}$$

Note that compared to the traditional Y-junction, where only the tapering width was optimized, inverse design methods allow to optimize also the material composition internal to the structure, suggesting those pixels where the material needs to be deposited. The function $\phi(x,y)$, then, is let evolving ether via an equation of motion or via gradients, that minimizes an objective function corresponding to the targeting functionality. Many optimization techniques can apply to find convenient trajectories in the enormous space of possible designs. While there is no guarantee that the solution found corresponds to a global minimum, it is anyway possible to find designs that perform remarkably well. As an example, an integrated polarization beam splitter has been realized, with internal pixels composed of silicon and air, on top of a silica substrate, [49]. The structure accounts for 20x20 pixels with sides of 120 nm, for a remarkably low device footprint of $2.4\mu m \times 2.4\mu m$. The device is able to discriminate between TE and TM input light polarizations and transmit them accordingly in one of the two output channels, with TE and TM transmission efficiencies of 71% and 80%, and corresponding extinc-

tion ratios of 11.8 dB and 11.1 dB, respectively. The same machinery can be applied for optimizing compact on-chip wavelength demultiplexers. In this application, the optical signal will propagate to a specific output channel depending on its wavelength [50].

2.6.4 Modulators

Within the brain, synaptic plasticity continuously adapts the weight of these connections both while wakefulness as well as during sleeping [51]. The scheme of an artificial neuron, inspired by biology, also considers connections whose weight values are trainable parameters. While these values are generally codified as real numbers in ANN models, they can be extended in photonics to the complex domain, thanks to the complex nature of light. Light modulators can be seen, under this perspective, as tools for applying complex weights directly on the chip, by changing both the amplitude and phase of the propagating optical signal.

Modulators are essential building blocks in silicon photonics, as phase and intensity are two key quantities that can encode information optically. The important characteristics of modulators include extinction ratio, insertion losses, modulation speed, energy consumption, footprint area, and modulation efficiency, defined as the product of the π-phase-shift voltage and modulator length ($V_\pi L$). These properties vary between phase and amplitude modulators.

Phase modulation To externally tune the phase of an optical signal, a possible strategy is to modulate the real part of the refractive index along the waveguide where the optical mode propagates. In silicon, the index tuning mechanisms are based on the thermo-optic effects or on the free-carrier absorption and free carrier dispersion (also known as plasma dispersion).

Silicon waveguides benefit of a relatively large thermo-optic coefficient $dn/dT = 1.86 \times 10^{-4} K^{-1}$, measured at room temperature $T = 300K$ [52], that favors the refractive index tuning. Thermal heating in a waveguide can be achieved by positioning a metal filament in the proximity of the waveguide during fabrication, or by doping the waveguide of donor carriers. Then a current is injected into the metal core or the doped waveguide which, by Joule effect, causes the heating of the resistor. Heat flow to the waveguide changes its temperature which in turn changes its refractive index. Thermal heating works at the KHz regime, which may result slow for applications

requiring faster modulation. Moreover, the heat can be transmitted to other structures across the chip, creating undesirable thermal cross talks. Nevertheless, photonics weights that exploit thermal heaters have been proposed. An example is reported in [53], where a photonic perceptron relies on the trainable complex modulation provided by thermal heaters.

An alternative way to encode data in the real part of the refractive index of a waveguide is to control the free carriers concentration in the waveguide, which in turn allows for free carrier dispersion and free carrier absorption modulations. The carrier concentration can be controlled by lateral p-n doping the waveguide and thermal heating.

When operating in a reversed-bias condition, the p-n junction depletes free carriers from the waveguide, in amounts proportional to the voltage applied. Free carrier dispersion induces thus phase shifts in the optical propagating light, which may serve in the realization of a complex weight in neural network architectures. Due to the relatively weak free carrier dispersion coefficient of silicon [54], carrier-depletion effects require long phase shift regions (typically millimeters), while maintaining low drive voltages to the p-n junction. For example, a phase shifter with $V_\pi L = 1.47$ V/cm, at a p-n junction reversed bias of -0.5 V is adopted in [55].

On the other hand, when forward biasing the p-n junction, free carriers are injected into the waveguide. As a result, free carrier dispersion is now accompanied by extra losses induced by free carrier absorption. Modulators based on carrier injection, still modulate the phase of the propagating light, benefiting of lower dimensions than those based on carrier depletion, but exhibit a lower bandwidth and higher losses. As a result, there is a trade-off between optical losses, footprint, and the $V_\pi L$ voltage, that needs to be found for the specific application at hand.

Amplitude modulation The modulation of the optical signal amplitude is especially required in optical neural networks that rely on incoherent light, where phase encoding does not produce any effects. An amplitude modulation can be achieved by placing phase modulators, like the ones just described, within an interferometric structure. When thermal or electrical phase modulators are placed along an arm of a Mach Zehnder interferometer, for example, they allow for different interference conditions at the output recombination of the two arms, thus changing the amplitude of the output signal. An example of a silicon modulator optimized for 56 Gb/s operation,

relying on carrier depletion over mm regions, is reported in [55]. In this application, a slower thermal phase shifter is also engaged, to provide a correct quadrature bias for the Mach Zehnder interferometer. In [56] instead, the authors combine amplitude and phase modulators by cascading 56 programmable Mach Zehnder interferometers, to realize an integrated photonic circuit for deep learning applications.

Alternatively, hybrid waveguides can be engineered to compensate for the lack of electro-optical effect in silicon. For example, other materials with favorable index modulation properties than silicon can be realized in the proximity of the silicon core, and hence perturb the propagation signal. Lithium niobate [57] and graphene [58] hybrid integration are an example.

Lately, modulators gated by phase change materials have been also proposed [59]. Phase change materials have the ability to reversibly shift from an amorphous to a crystalline phase and, when placed on top of a waveguide, attenuate diversely the propagating mode, thus applying a modulation. These materials can be addressed both electronically, or by dedicated optical pulses that heat the material and induce the phase transition. In the context of neural networks, phase change materials represent promising weight candidates, especially during inference tasks, when the optimal weights are kept fixed. If properly codified in the state of different segments of phase change materials, no additional energy is required [60], thus providing advantages in terms of energy efficiency with respect to the electronic/thermal counterparts [61].

2.7 Silicon nonlinear optics

Nonlinear optics in silicon photonics is desirable for neuromorphic applications, since the importance of nonlinear transformations in ANNs. The study of nonlinear optics was accelerated when in the '60s, lasers begin appearing on the commercial market, paving the way for the study of nonlinear phenomena in materials with weak dielectric susceptibility.

The nonlinear optical properties of a material are described by the relation between the macroscopic polarization P, and an externally applied field E [62]:

$$P = \epsilon_0(\chi^{(1)}E + \chi^{(2)}E^2 + \chi^{(3)}E^3 + \dots), \tag{2.15}$$

with ϵ_0 indicating the vacuum permittivity and $\chi^{(i)}$ the dielectric susceptibilities of the material. For low exciting optical power, the relation involves

only $\chi^{(1)}$ and is linear. Higher-order terms become important as the value of the external electric field E approaches the one binding electrons and nucleus in the atom, typically $10^5 - 10^8 V/m$ [36].

To better understand the role of nonlinear effects, Maxwell equations can be manipulated in the simple case of a homogeneous and isotropic dielectric medium, in terms of the polarization $P = \epsilon_0 \chi E + P_{NL}$, in the following way [36]:

$$\nabla^2 E - \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} = -\mu_0 \frac{\partial^2 P_{NL}}{\partial t^2}, \tag{2.16}$$

where the electric field E must satisfy an equation where the nonlinear contribution of the polarization P_{NL} acts as a radiative source inside the material, and this polarization term is itself depending on the input electric field E due to Eq. 2.15.

In center-symmetric materials, like silicon, all the susceptibilities of even order vanish due to symmetry considerations [36]. The third order susceptibility, therefore, becomes the lowest effective order, with the real and imaginary part that amounts to $\chi^{(3)}_{Im} = 10^{-19}\frac{m^2}{V^2}$ and $\chi^{(3)}_{Re} = 10^{-5}\frac{m^2}{V^2}$, respectively. The real part is responsible for the Kerr effect while the imaginary part is linked to an additional absorption process inside the material named Two-Photon Absorption (TPA), both occurring in timescales of the order of femtoseconds. The Kerr effect introduces a change in the real part of the refractive index which is proportional to the intensity of the electric field (I):

$$\Delta n_{kerr} = n_2 I \tag{2.17}$$

with n_2 the Kerr coefficient, given by:

$$n_2 = \frac{3\eta_0}{n^2 \epsilon_0}\chi^{(3)}, \tag{2.18}$$

where $\eta_0 = \sqrt{\mu_0/\epsilon_0}$ is the impedence of the medium, being μ_0 the vacuum magnetic permeability. The optical Kerr effect is a self-induced effect where the phase velocity of the wave depends on the wave's own intensity. In silicon, at $\lambda \sim 1550nm$, the Kerr coefficient is $n_2 \sim 0.45 \times 10^{-17} m^2/W$ [62]. The optical Kerr effect is responsible for effects like self-phase modulation, self-focusing, and spatial solitons [36]. TPA indicates the simultaneous absorption of two photons, with equal or different frequencies, by electrons that move from the valence band to the conduction band of the semiconductor. In silicon for example, at a wavelength of $1.55\mu m$, the energy of a single

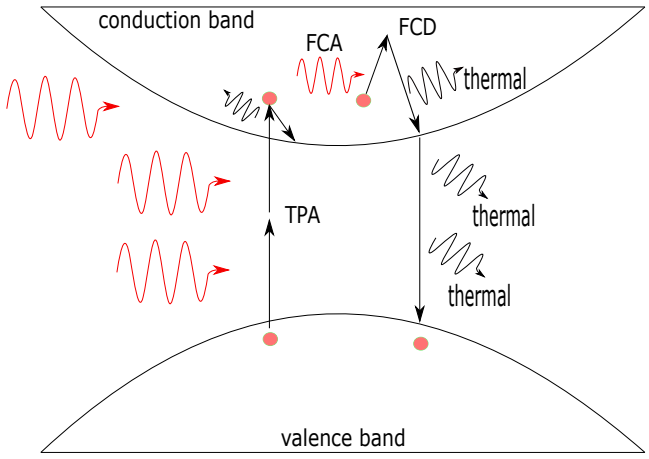

Figure 2.14: TPA allows the promotion of valence electrons to the conduction band in silicon, at communication wavelengths. Extra carriers in the conduction band lead to FCD and to FCA. The process also involves phonons release with consequent material heating.

photon $h\nu \sim 0.81eV$ is lower than the silicon bandgap ($1.1eV$) and it is not sufficient to excite electrons in the conduction band and holes in the valence band. On the other hand, two photons can induce this excitation via TPA. Fig. 2.14 describes the chain of events that TPA induces in silicon: valence electrons are excited in the conduction band, increasing the number of free carriers and leading to variations in the refractive index known as free carrier dispersion (FCD). Electrons in the conduction band in turn absorb the propagating light by free carrier absorption (FCA) and release the energy by phonons emission that heats the material. Moreover, since silicon is a semiconductor with an indirect bandgap, the electron decay from the conduction to the valence band is accompanied by phononic emission too.

2.7.1 Thermal nonlinearity

The intra-band relaxation of free carriers occurs by emitting phonons and consequently heating the silicon waveguide. The refractive index variation can be estimated by:

$$n_0(T, w) \approx n_0(T_0, w) + \left.\frac{dn_0}{dT}\right|_{T_0} (T - T_0) \tag{2.19}$$

with $\frac{dn_0}{dT} = 1.86 \times 10^{-4} k^{-1}$ the thermo optic coefficient evaluated at room temperature $T_0 = 300K$, and $n_0(T_0, w) \approx 3.5$ at $\lambda = 1.55\mu m$.
When working at steady state $\Delta T = T - T_0$ can be related to the field intensity by the following relation :

$$\Delta T = \gamma_{lin}|E|^2 + \gamma_{TPA}|E|^4 + \gamma_{FCA}|E|^6 \tag{2.20}$$

where γ_{lin}, γ_{TPA}, and γ_{FCA} are coefficients that refer to the linear, TPA, and FCA absorption mechanisms respectively. Thus, ΔT is linear with light intensity and so is Δn, until TPA and FCA come into play:

$$n_0(T, w) \approx n_0(T_0, w) + \left.\frac{dn_0}{dT}\right|_{T_0} (\gamma_{lin}|E|^2 + \gamma_{TPA}|E|^4 + \gamma_{FCA}|E|^6) \tag{2.21}$$

The temperature dynamics of a silicon structure, when absorbing optical power, is typically described according to the Fourier law:

$$\frac{d\Delta T(t)}{dt} = -\frac{\Delta T(t)}{\tau_{th}} + \frac{P_{abs}}{mC_p}, \tag{2.22}$$

with P_{abs} indicating the absorbed power by the material that causes heating, m the mass, C_p the silicon specific heat at constant pressure, and τ_{th} the thermal lifetime. The thermal lifetime is the typical time needed by the waveguide to relax to room temperature T_0 in absence of optical excitation. It includes the heat loss in the surrounding medium. Thermal effects occur on timescales down to tens of nanoseconds for micron-wide waveguide [63]. Nevertheless, these timescales strictly depend also on the fabrication process, and it was recently reported that a more sophisticated equation than the Fourier low may be necessary for a correct description [64].

2.7.2 Free carriers nonlinearity

The variation of the free carrier density in the conduction band alters both the real (Δn^{FC}) and imaginary ($\Delta \alpha^{FC}$) parts of the refractive index of the material. Specifically [65][66]:

$$\Delta n^{FC} = \left.\frac{dn}{dN}\right|_{N_0} \Delta N \overset{Si}{\simeq} -1.73 \times 10^{-21} \cdot \Delta N \tag{2.23}$$

$$\Delta\alpha^{FC} = \left.\frac{d\alpha}{dN}\right|_{N_0} \Delta N \overset{Si}{\simeq} 1.45 \times 10^{-17} \cdot \Delta N \tag{2.24}$$

where $\Delta N = N - N_0$ indicates the variation in free carrier density population concerning the equilibrium value. Its dynamical behaviour is described by the following rate equation:

$$\frac{d\Delta N}{dt} = -\gamma_{FC}\Delta N + G, \tag{2.25}$$

where γ_{FC} is a phenomenological quantity taking into account the recombination rates of electrons from the conduction to the valence band and G is the total generation rate per unit volume. At optical communication wavelengths, free carriers are mainly generated through TPA and thus G can be written as:

$$G = \frac{P_{TPA}}{2\hbar w V}, \tag{2.26}$$

where P_{TPA} is the TPA absorbed power. Nonlinearities due to carrier effects occur on timescales ranging from a few to tens of ns [64] [67]. This free carrier lifetime can be further reduced when embedding the waveguide in a p-n junction that, when operated in reverse biased, speeds up the depletion of the generated carriers from the device, down to tens of picoseconds [68].

Thermal and free carrier variations are responsible for nonlinear effects in silicon only when generated from the propagating optical power. On the other hand, if their variations are produced by means of thermal heaters and p-n junctions, for instance, the effects are only visible in the linear susceptibility $\Delta\chi^{(1)}$. The excess polarization $\Delta\mathbf{P} = \epsilon_0\Delta\chi^{(1)}\mathbf{E}$, maintains indeed a linear relationship with the input optical field $\mathbf{E}$, although the temperature and free carrier variations.

An interesting way to trigger nonlinear effects in silicon photonics while using relatively low input optical power is through resonance structures. In the next chapter, silicon-based microring resonator structures are presented, and accompanied by discussing its usefulness in a wide variety of applications in neuromorphic photonics.

Chapter 3

Silicon microring resonators

An optical microring resonator consists of a bent waveguide closed on itself. A microresonator placed on an optical chip represents an isolated system unless it is coupled via an evanescent field overlap with other waveguides in its close proximity. Generally, microring resonators are designed in the two configurations illustrated in Fig. 3.1. An input bus waveguide provides the incoming optical signal that is coupled to the microring, according to the coupling coefficient k_1, and transmitted to the through port, according to the coefficient t_1. A resonance condition inside the microring is satisfied when, after a round trip, light constructively interferes with itself, i.e. when the length of the ring's circumference is an entire number of wavelengths:

$$\lambda_m m = 2\pi r \quad ; \quad \beta_m 2\pi r = m2\pi \tag{3.1}$$

where r is the microring radius and m is the azimuthal mode number. In terms of the vacuum wavelength λ_0, Eq. becomes:

$$\lambda_{0,m} = \frac{n_{eff}(\lambda)2\pi r}{m} \tag{3.2}$$

The microring is in the all-pass configuration when it is only coupled to one input waveguide. In this case, all wavelengths can be transmitted from the input to the through port. On the other hand, if an additional output waveguide is coupled to the microring, the latter is in the add-drop configuration. The output waveguide offers the possibility to extract the resonant wavelengths from the microring and eventually mix them to an incoming optical signal propagating from the add port. While the all-pass configuration may serve as a frequency-dependent delay line, a notch filter or as a phase shifter,

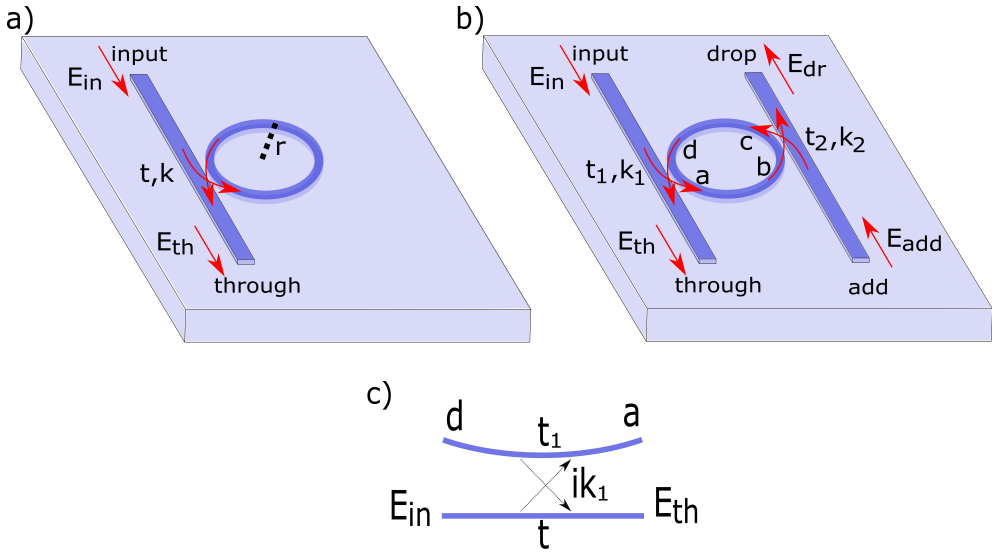

Figure 3.1: The two main configurations of microring resonators. a) The all-pass and b) the add-drop filter configurations. c) Coupling region between the input bus waveguide and the microring waveguide.

in the add-drop configuration the microring acts like a pass-band filter and a wavelength router, thus resulting in a critical component for optical networks on chip, especially for telecom applications [31].

3.1 Linear regime

3.1.1 Frequency response

The steady-state frequency response of a microring working in the add-drop configuration and in a linear regime can be estimated through a scattering matrix approach [36]. Here, the optical power exchanged via the evanescent fields between the microring and the waveguides is described by two directional couplers characterized by the parameters k_i, t_i, with $i = 1, 2$. The directional coupler is assumed to satisfy reciprocity and to be lossless, by imposing $k_i^2 + t_i^2 = 1$. This approximation is a better approximation the more the coupling region can be considered point-like. If for example, the coupling region is extended by elongating the portion of microring coupled to the external waveguides, propagation losses in that case should be considered.

The 2×2 directional coupler at the input waveguide can be described in a scattering matrix approach in the following way, referring to Fig. 3.1(b)(c) for the a, b, c and d fields:

$$\begin{pmatrix} a \\ E_{th} \end{pmatrix} = \begin{pmatrix} ik_1 & t_1 \\ t_1 & ik_1 \end{pmatrix} \begin{pmatrix} E_{in} \\ d \end{pmatrix}. \tag{3.3}$$

In the same way, the directional coupler at the output waveguide results:

$$\begin{pmatrix} E_{dr} \\ c \end{pmatrix} = \begin{pmatrix} ik_2 & t_2 \\ t_2 & ik_2 \end{pmatrix} \begin{pmatrix} b \\ E_{ad} \end{pmatrix}. \tag{3.4}$$

The propagation inside the resonator is described by means of the roundtrip *field* attenuation factor $a_{rt} = e^{\frac{-\alpha p}{2}}$, with p the microring perimeter and α the microring losses per unit length, and the roundtrip phase accumulated $\phi_{rt} = \beta p$, with β the propagation constant. For example:

$$d = c(\sqrt{a_{rt}} e^{-i\phi_{rt}/2}) = \cdots = t_2 a_{rt} e^{-i\phi_{rt}} a, \tag{3.5}$$

By expressing a as a function of d, as in Eq. 3.3, and substituting Eq. 3.5, one obtain the following expression for the internal field a:

$$a = \frac{ik_1 E_{in}}{1 - t_1 t_2 a_{rt} e^{-i\phi_{rt}}}, \tag{3.6}$$

from which also the fields $b = \sqrt{a_{rt}} e^{-i\phi_{rt}/2} a$ and the field $c = t_2 \sqrt{a_{rt}} e^{-i\phi_{rt}/2} a$. As can be noted, the power within the microresonator is equally distributed ($|a| \approx |b| \approx |c| \approx |d|$) whenever the internal losses and coupling coefficients are small, i.e. when $a_{rt} \approx 1$ and $t_1, t_2 \approx 1$.

Substituting the expression of a of Eq. 3.6, in Eq 3.3 and Eq. 3.4, the through (T) and drop (D) optical transmissions are finally extracted [69]:

$$T = \left| \frac{E_{th}}{E_{in}} \right|^2 = \left| \frac{t_1 - t_2 a_{rt} e^{-i\phi_{rt}}}{1 - t_1 t_2 a_{rt} e^{-i\phi_{rt}}} \right|^2 \tag{3.7}$$

$$D = \left| \frac{E_{dr}}{E_{in}} \right|^2 = \left| \frac{-k_1 k_2 \sqrt{a_{rt}} e^{-i\phi_{rt}/2}}{1 - t_1 t_2 a_{rt} e^{-i\phi_{rt}}} \right|^2 \tag{3.8}$$

Note that the equations of microring in all-pass configuration can be obtained from the precedent one, by simply using $t_2 = 1$ ($k_2 = 0$).

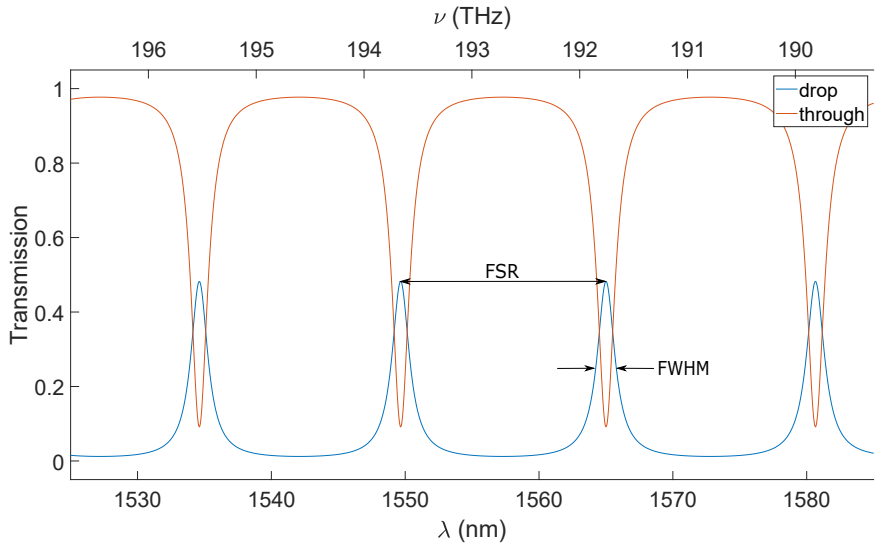

Figure 3.2: Through and drop transmissions of a microring resonator in add-drop filter configuration, having radius $r = 6.75um$, $k^2 = 0.2$ and $a_{rt} = 0.9$.

The drop transmission signal of a microring in an add-drop filter configuration is shown in Fig. 3.2, as a function of the input wavelength. The spectra are characterized by a series of Lorentzian peaks at the drop port (blue curve) and a series of Lorentzian dips at the through port (red curve), expressing the lack of optical power at the through port whenever the wavelength is resonant with a microring resonance.

Several figures of merit can be defined to describe a microresonator. The distance between two consecutive resonances is defined as the Free Spectral Range (FSR). It is obtained by a first-order Taylor expansion of the propagation constant $\beta = n_{eff}k_0 = n_{eff}2\pi/\lambda$ respect to λ and noting that, from Eq. 3.1, $\beta_{m-1} - \beta_m = -\frac{1}{R}$. The FSR results:

$$FSR = \frac{\lambda_m^2}{n_g(\lambda_m)2\pi R} \tag{3.9}$$

The average spectral distance between two resonances thus decreases as the resonator increases in size.

Another important feature is the Full Width at Half Maximum (FWHM) of the Lorentzian resonances, which is related to the internal losses and to

the coupling coefficient according to [70]:

$$FWHM = \frac{(1 - t_1 t_2 a_{rt})\lambda_m^2}{\pi n_g(\lambda_m) p \sqrt{t_1 t_2 a_{rt}}} \tag{3.10}$$

The quality factor (Q) is defined as the ratio between the m^{th} order resonant wavelength and its FWHM, and can be expressed as:

$$Q = \frac{\pi n_g(\lambda_m) p t_1 t_2 \sqrt{a_{rt}}}{(1 - t_1 t_2 a_{rt})\lambda_m} \tag{3.11}$$

As Eq. 3.11 shows, high Q-microrings are characterized by weak coupling with the external waveguides, low internal losses and high radius (perimeters).

Finally, the enhancement factor (EF) refers to the ratio between the field inside the microresonator and the one at the input port.

$$EF = \frac{ik_1}{1 - t_1 t_2 a_{rt} e^{-i\phi_{rt}}} \tag{3.12}$$

As Eq. 3.12 shows, a high enhancement factor is achieved by reducing the internal losses ($a_{rt} \rightarrow 1$) and by reducing the coupling coefficients. For example, $EF \sim 1/k$ when considering a resonant wavelength, a symmetric coupling $k_1 = k_2$, and low internal losses $a_{rt} \approx 1$. In this condition, a coupling coefficient $k = 0.01$, allows for optical power within the microring hundred times larger than the input one. Microring resonators are for this reason useful tools to reduce power requirements in nonlinear optics applications. The spectrum in Fig. 3.2 is described for example by a $FSR = 15.2nm$, a $FWHM = 1.55nm$, a $Q = 0.7$ and an $EF = 1.78$ calculated at resonance.

Note that the values plotted in Fig. 3.2 represent the stationary optical power achieved at the drop and through ports of the microring, in a linear regime. The evolution of the system towards those stationary values is a consequence of the time the microring takes to load the optical field within the cavity, up to the steady state. This time is referred to as photon lifetime (τ_{ph}), and is linked to the quality factor through:

$$\tau_{ph} = \frac{Q}{2\pi\nu}. \tag{3.13}$$

An input signal that is modulated in time faster than the photon lifetime, for example, continuously keeps the microring internal field in a transient

state, which in turn will be observed at the through and drop ports. These transients will be investigated and exploited in section 4.2.3, for a time-delay RC application involving a microring resonator.

3.1.2 Linear time response

The linear time response of a microring resonator can be calculated from Eq. 3.6, by firstly expanding it around a resonance frequency w_0 and then applying an inverse Fourier transform [66]. Let us assume low propagation losses within the microring ($\alpha p << 1$) and weak coupling coefficients ($k_{1,2} << 1$), then:

$$\frac{d\tilde{U}(t)}{dt} = \left[-i(w_p - w_0) - \left(\frac{1}{\tau_{k1}} + \frac{1}{\tau_{k2}} + \frac{1}{\tau_0}\right)\right]\tilde{U}(t) + i\sqrt{\frac{2}{\tau_{k1}}}E_{in}, \qquad (3.14)$$

where $\tilde{U}(t)$ represents the slowly varying envelope of the internal energy amplitude $U(t) = \tilde{U}(t)e^{i(w_p)}$, $w_p = 2\pi/\lambda_p$ the frequency of the input exciting field $E_{in}(t) = \tilde{E_{in}}(t)e^{iw_p t}$, and $1/\tau_{k1}$, $1/\tau_{k2}$ and $1/\tau_0$ the decay rates of the internal energy induced by the two coupling regions and the linear losses within the microring resonator, respectively. As can be noted, Eq. 3.14 describes the energy amplitude within the microring as a mass-spring system subject to a specific driving force. When no external forces are applied, the energy within the microring decays with a total rate $1/\tau_{tot} = \left(\frac{1}{\tau_{k1}} + \frac{1}{\tau_{k2}} + \frac{1}{\tau_0}\right)$. On the other hand, when the driving force is present, the cavity exhibits a resonance when $w_p = w_0$.

3.2 Nonlinear time response

The possibility to enhance the optical field within microring resonators allows for nonlinear effects in silicon photonics. As reported in section 2.7, two-photon absorption leads to the generation of extra free carriers in the conduction band, which in turn increases the optical losses and the temperature within the microring waveguide, changing as a result also the refractive index of the silicon waveguide. For a microring resonator, these effects can be taken into account in Eq. 3.14 by additional parameters. Refractive index variations modify here the resonance wavelength positions (Δw_{nl}^R) and introduce extra losses that increase the decay rate of the energy field within

the microring (Δw_{nl}^{I}). It is possible to account for these nonlinear effects, by rewriting Eq. 3.14 as:

$$\frac{d\tilde{U}(t)}{dt} = \left[-i\left(w_p - \left(w_0 - \Delta w_{nl}^{R}(t)\right)\right) - \left(\frac{1}{\tau_{tot}} + \Delta w_{nl}^{I}(t)\right)\right]\tilde{U}(t) + i\sqrt{\frac{2}{\tau_{k1}}}E_{in}. \tag{3.15}$$

The microring resonance is thus modified with respect to its cold position (w_0) depending on the nonlinear effects taking place within the microring. The nonlinear resonance shift can be expressed as a contribution of a thermal, free carrier and a Kerr shifts:

$$\begin{aligned}\frac{\Delta w_{nl}}{w_0} &= 1/w_0\left(\Delta w_{TOE} + \Delta w_{FC} + \Delta w_{kerr}\right) \\ &= \frac{\Gamma_c}{n_{Si}}\frac{dn}{dT}\Delta T + \frac{\Gamma_c}{n_{Si}}\frac{dn}{dN}\Delta N + \frac{cn_2}{n_{Si}^2 V_{eff}}\tilde{U}(t)^2,\end{aligned} \tag{3.16}$$

with Γ_c the confinement factor, n_{Si} the silicon refractive index, n_2 the Kerr coefficient, dn/dT the thermo-optic coefficient at room temperature, dn/dN the free carrier dispersion coefficient, ΔT the mode-averaged temperature difference between the circular waveguide of the microring and its surroundings, ΔN the excess carrier concentration, V_{eff} the mode effective volume. The integration of Eq. 3.16 requires the knowledge of the temperature and free carrier dynamics within the waveguide, as given by Eq. 2.22 and Eq. 2.25. All combined, they provide a set of three coupled nonlinear differential equations that fully describe the nonlinear cavity dynamics.

3.3 Photonics computing with microring resonators

Both the linear and nonlinear response of microring resonators has been investigated in the context of photonics computing. While the microring linear frequency response suites the realization of photonic weights, its nonlinear response is proposed as a possible nonlinearity candidate for photonic neurons. The aim of this section is to provide an overview of these applications.

3.3.1 Linear response as weight

An important component of an artificial (and also biological) neuron are the input weights. The weights are trainable parameters that adapt and allow the

network to solve specific tasks. The linear filtering capabilities of microring resonators in add-drop filter configuration make them very attractive components for the realization of input weights. A microring resonator placed at the input port of a photonic neuron is indeed able to transmit only a portion of the incoming signal, according to Eq. 3.7 and 3.8, depending on the detuning between the input and the resonance wavelengths. Whenever the detuning is externally controlled by shifting the microring resonance wavelength, thus changing its transmission properties, an analog and trainable weight can be obtained by means of a microring. The idea is particularly suitable when the information to process is encoded onto several multiplexed wavelength channels. In this case, it is natural to deliver the weight operation to an extended set of microring resonators, the weight bank, still placed at the input of a photonic neuron. In doing so, each wavelength channel can be uniquely weighted, and the intrinsic parallelism of photonics hardware is exploited. In a common weight bank geometry (see Fig. 3.3), a set of microring resonators share the same input and output bus waveguides. Multiplexed wavelength signals enter the input port, are filtered by the weight bank, and their resulting two weighted versions are transmitted to the through and drop ports, where a balanced photodiode finalizes the incoherent optical sum. The generated photocurrent encodes in this way the input weighted sum of the photonic neuron. To note that the use of a balanced photodiode allows to apply both positive and negative weights in the microring weight bank. Indeed, the weight will be zero whenever the microring equally splits the energy between the two ports, and will be positive or negative whenever an unbalance is present [71].

The microring extinction ratio sets the range of weight values that can be applied. Ideally, the goal is to obtain the maximum extinction ratio, by having a complete energy transfer to the drop port when the incoming wavelength is in resonance, or a complete energy transfer to the through the port when it is out of resonance. Designing the microring in critical coupling allows to best approach this situation [72]. Equally important is the radius r of the microrings, as linked to their resonance wavelengths ($\lambda = n_{eff} 2\pi r/m$, with $m = 1, 2, 3, ...$). Designing a weight bank with microrings that have a slightly different radius, allows the separation of the individual resonance spectra. When the resonance spectra are completely separated, i.e. when there is no optical cross-talk, the device is described as an ensemble of independent microrings, each one applying a weight to a specific wavelength channel. Nevertheless, optical cross-talk between the microrings is also a

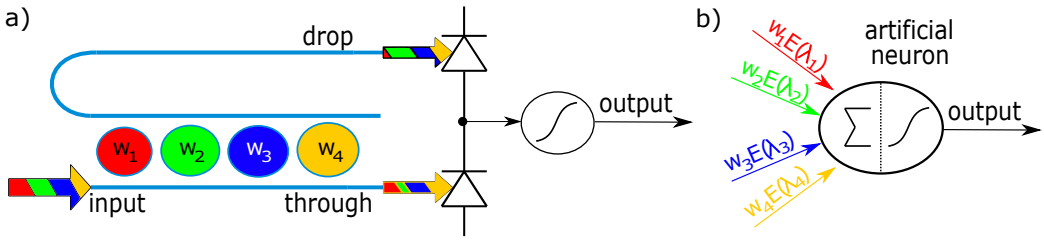

Figure 3.3: a) A weight bank composed of an array of microring resonators in an add-drop filter configuration. The structure is specifically designed to take advantage of wavelength-division-multiplexing (WDM) and applies specific weights (w_i) to the incoming wavelength channels (λ_i) by tuning the transmission of the microrings. A balanced photodetector integrates the through and drop weighted signals, and enables both positive and negative weights. b) Artificial/biological neuron representation, where input values are firstly weighted, then integrated and finally used as an argument for the neuron nonlinear activation function. Differently from a biological neuron where a multidimensional input requires several physical channels, photonics equipped by WDM can employ only one.

subject of recent studies [73].

The FSR is also affected by the radius according to Eq. 3.9, and represents a critical limiting factor for the number of wavelength channels that the weight-bank can access and weight.

Because of fabrication errors, every microring is realized slightly differently with respect to the values of the nominal parameters set in the design process. Additionally, environmental changes like temperature variations, and thermal and electrical cross-talks between the microrings also interfere with the weight bank. As a result, precise calibration and control strategies have been developed. The latest progress relies on appropriate sensing elements placed at each microring resonator site, to realize integrated feedback controls, aiming to set and maintain a desired weight [74]. For example, such a sensor can be realized by embedding a heater in the microring waveguide, by lightly doping [75]. In this case, when an electric current is applied, a variation in the applied voltage can be sensed whenever light circulates the microring. In fact, donor-induced extra losses in the waveguide will produce new electron-hole pairs, which in turn lower the conductance. Note that the same sensor can also be used to tune the weight of the microring, by heating

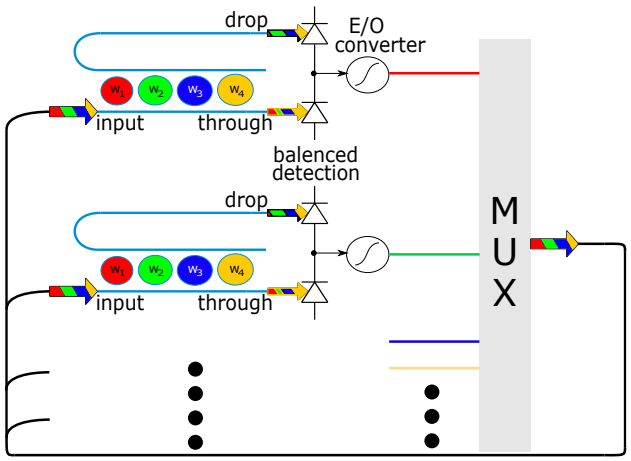

Figure 3.4: Photonics implementation of neural network adopting the broadcast-and-weight architecture [79]. A broadcast interconnect drives the multi-wavelength optical signal to all the photonics neurons. A microring-based weight bank applies the input weights to the incoming signal, producing photocurrent that an electro-optical (E/O) converter nonlinearly imprints on a specific wavelength carrier, which is finally multiplexed into the broadcast interconnect. Each neuron emits at a specific wavelength, provided that it couples to the weight banks to be later on weight. From [73]

and thus shifting its resonance wavelength, with a precision up to 7 bits [76], and up to 9 bits when the control extends to noise sources other than the microrings [77]. Thermal stabilization of the photonic chip is an additional solution to balance temperature variations in the environment.

A first proposal of an integrated photonic network implementing microring weight banks was suggested in [78], in what is known as the "broadcast and weight" protocol. The idea, schematically represented in Fig. 3.4, is to multiplex N different wavelengths onto a single common medium that brings the signal to the photonic neurons and at the same time collects their optical response. The information is thus continuously reused by the network. Each photonic neuron is equipped with an N-microring weight bank, which independently acts on each input wavelength channel, followed by an electro-optical converter. At this stage, the photocurrent generated by the balanced photodiode is nonlinearly imprinted on an output light signal at one specific wavelength channel, which is in turn finally multiplexed in the common medium. The electro-optical conversion can be implemented in sev-

eral ways. For example, the electrical signal can be delivered to a laser close to the lasing threshold to trigger an optical spike. Or it can be used to drive a phase shifter in a Mach Zehnder modulator fed by a CW laser signal [80]. In a recent implementation, the electro-optical conversion is realized through a further microring resonator equipped with a pn-junction [79], which represents one of the possibilities to take advantage of microring resonators as photonic nonlinear activation functions (section 3.3.2).

Microring-based weight banks are emerging as promising platforms also for photonics Tensor Processing Units (TPUs), where they can perform matrix multiplications and speed up those applications that rely on matrix algebra. Machine learning, and in particular deep neural networks, heavily depend on these operations. An example of photonic TPU employing microring-based weight banks is reported in [81]. A Matrix-Vector Multiplication (MVU) operation $O = XI$ is performed, being O and I respectively the output and input m-dimensional vectors, and X an $m \times m$ matrix. In the experimental implementation, X has dimensionality 4×4 and is realized via 4 weight banks, each one composed by an array of 4 microrings. The time-of-flight latency for performing an operation is here only limited by the detection speed, which can be as short as tens of picoseconds in modern photodetectors. Note also that in machine learning tasks, once concluded the training phase, the weights are usually fixed or slowly updated in time, with respect to the input vector. In this situation, the device becomes particularly efficient, as once the weights are loaded by tuning the heaters of each microring, the structure is able to process a fast input signal requiring only a constant energy power to keep the microring transmissions at the trained value. Additionally, the energy efficiency can be further improved by adopting phase change materials (PCMs), as they allow for nonvolatile photonics memory that only demands energy when the weight needs to be updated. This solution has been investigated with PCMs based on $Ge_2Sb_2Se_2$, $30nm$ thin and $250nm$ wide, placed on top of the waveguide and arranged in a grating fashion, within a structure closely similar to the microring-based weight bank previously discussed [82]. By local electrostatic heating, the PCMs can be individually and reversibly switched between an amorphous and a crystalline phase, characterized by different absorption coefficients. The authors numerically showed that by cascading 15 PCMs a memory with total insertion losses limited to only $1dB$ (all PCMs in the amorphous state), extinction ratio up to $3.5dB$ (all PCMs in the crystalline state) and 4-bit resolution can be achieved.

3.3.2 Nonlinear response as activation function

In ANNs the original information is mapped into a higher-dimensional space and processed by nonlinear functions. As mentioned in section 2.1, each artificial neuron initially integrates the signal arriving from the input channels and then applies a nonlinear activation function. Thanks to their nonlinear behaviors, microring resonators have been proposed as photonic candidates for applying the nonlinear activation function of a neuron. By looking back at Fig. 2.2, sigmoid and ReLU activation functions are quite commonly used by the computer science community [8][83]. These activation functions can be realized by microring resonators. Furthermore, spiking neural networks rely on spiking emissions from the neurons, as explained in section 2.4. The microring can also replicate the spiking behavior [84][65][85].

Steady state nonlinearity

The mechanism exploited to achieve a sigmoid-like activation function is the microring thermal bistable behavior, schematically represented in Fig. 3.5(a). The thermal bistability occurs whenever the thermal nonlinear effects overcome the free carrier ones, and the pump wavelength λ_p is initially slightly red-shifted with respect to the cold resonance λ_0 ($\lambda_p > \lambda_0$) [86]. As a result, while increasing the input optical power, the microring internal optical energy is firstly increased linearly, until nonlinear phenomena triggered by TPA start to occur. The microring resonance position is, hence, red-shifted due to the predominance of the thermal nonlinearities, moving closer to λ_p (refer to Fig. 3.5(b)), which in turn favors a higher field in the cavity. A positive feedback mechanism is finally triggered, ending up with $\lambda_p < \lambda_0$ and high optical power at the drop port. The effects of larger input power are now hindered since the resonance is further red-shifted, moving away from λ_p (optical limiting). At this point, if the optical power reverts and starts to decrease, the microring internal optical power decreases as well, firstly hindered by the resonance blue-shift (cooldown), which now moves towards λ_p. Then, it jumps to a lower level thanks to a similar negative feedback mechanism. The above considerations translate in the two curves of Fig. 3.5(a): the red curve indicates the step when the input power is increased ("hot" curve) while the blue curve indicates the step when it is decreased ("cold" curve).

From a neuromorphic processing point of view, despite the microring

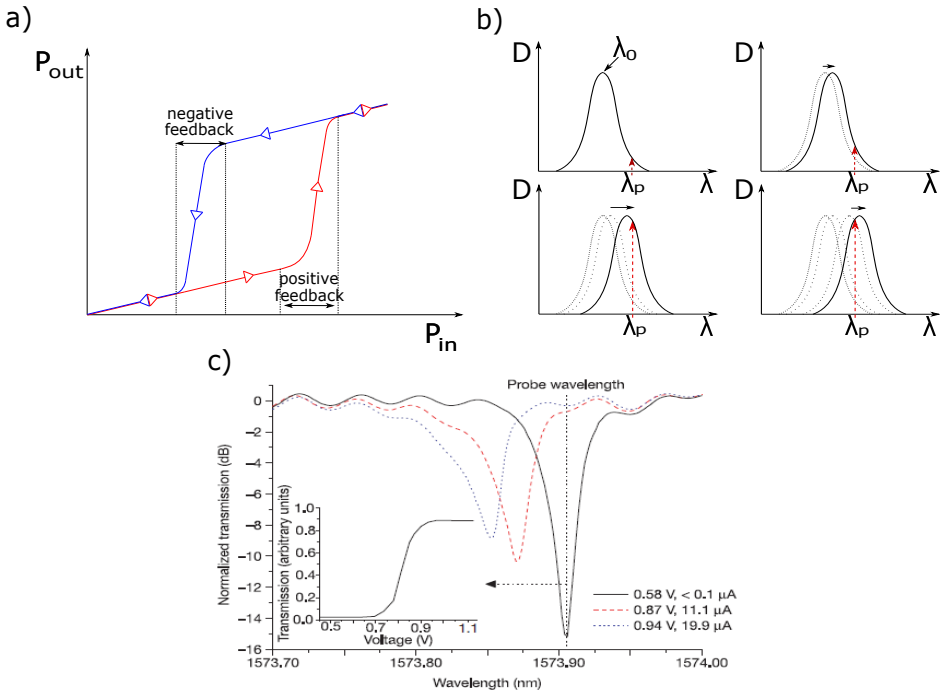

Figure 3.5: a) Thermal bistability of an add-drop microring resonator similar to a sigmoid activation function. The red curve represents the "hot" curve, obtained by increasing the input power while the blue curve indicates the "cold" curve obtained by decreasing the input power. P_{in} and P_{out} represent the input and drop optical power. b) Resonance shift position along the hot curve, for different P_{in}. c) Example of ReLU activation function obtained by embedding a microring (in all-pass configuration) in a p-n junction and varying the input voltage.

"hot" curve resembles to a sigmoid activation function, photonics implementation needs to account also for the "cold" curve. Indeed, the microring that has been promoted to the higher state of the bistability loop while processing the optical bit x_i, can not exhibit the same 'jump' for the next bit x_{i+1}, as it will respond through the 'cold' curve in Fig. 3.5. This is a remarkable difference with respect to the simpler sigmoid function adopted in ANNs. Eventually, it can be avoided by separating the input optical bits by a sufficient time for the microring to cool down. This happens at the cost of a reduced processing speed. Note, however, that the thermal bistability can be exploited as a memory-based nonlinearity, resulting in an attractive ingredi-

ent for photonic neural networks. An example of an all-on-chip feed-forward network that exploits the thermal nonlinearity of three microring resonators is proposed in [87].

Microring nonlinearities are not limited to TPA-triggered processes. Another way to obtain a steady state nonlinear transformation is with the use of a p-n junction in a microring. In this case, the p-n junction operates in forward bias condition and modifies the free carrier concentration within the microring waveguide, hence changing the resonance position λ_0. As a result, the optical signal at the microring input port is transmitted differently, according to the voltage driving the p-n junction (refer to Fig. 3.5(c)). At this point, the device behaves just as a tuned filter, where the output power, either at the through or drop ports, still remains linear with respect to the input optical power. The method becomes nonlinear when the driving voltage to the p-n junction represents certain information to process. This (electrical) information is thus nonlinearly imprinted onto the optical signal propagated through the microring, via the correspondent transmission changes. The Lorentzian microring resonance describes the nonlinear function of this electro/optic nonlinear process. Fig.3.5(c) shows a ReLU-like activation function obtained with a microring in all pass configuration, in a starting configuration where $\lambda_p = \lambda_0$, when zero voltage is applied to the p-n junction. Consequently, the optical power at the through port is initially at the minimum. A larger voltage to the p-n junction drives away the resonance, increasing the transmitted optical power accordingly to the through port.

Photonic neurons equipped with this kind of nonlinearity have been realized in [88], within an end-to-end fully integrated feed-forward network. The network is designed for image classification tasks, with an input layer composed of a 5x6 matrix of grating couplers (150 μm x 140 μm). A laser beam from the top of the chip is opportunely collimated (beam diameter 870 μm), and shined on a specific letter designed on a custom-made Plexiglas holding frame. The transmitted light is thus collected by the grating couplers (41dB estimated losses) and routed to the first of three layers composing the network by means of nanophotonics waveguides, y-junction splitters, and waveguide crossings. The input light signals to each photonic neuron are carried by independent P-doped-intrinsic-N-doped current-controlled waveguide channels (500 μm long), that, when in forward bias, also allow to apply the trained weight. These signals are independently detected by silicon-germanium photodiodes (PD) and combined to generate an electronic input weighted sum.

This last is then amplified and converted to a voltage by a transimpedance amplifier (TIA) and finally drives the forward bias pn-junction of a microring resonator. Interestingly, all the layers in the network are externally powered by CW light (where the microring nonlinearity applies), mitigating scalability issues induced by optical losses.

P-n junction-equipped microrings also serve as neuron activation functions in [79]. The authors used in this case the broadcast-and-weight protocol to realize a feed-forward neural network model, having two hidden layers with two and eight neurons, respectively. The architecture has proven to be successful for fiber nonlinear compensation tasks, and fiber transmission signal recovery.

Nonlinear transients

Nonlinear transient dynamics are initiated whenever the time-varying input is faster than the system nonlinearity timescale of interest and are the base of the processing paradigm introduced with liquid state machine [3]. Transient dynamics based on both thermal and free carrier nonlinearity can be achieved. An example is reported in [89], where a reservoir computer made with a 4x4 matrix of silicon microring resonators, connected in a swirl topology, is numerically investigated. In this implementation, the microrings represent the physical nonlinear nodes of the network, while the connectivity matrix is given by splitting ratios, losses, and random phase shifts. The network is tested successfully on the 1-bit delayed XOR task, a nonlinear boolean operation that requires storing 1 bit of past information, and a proper nonlinearity transformation. Other implementations of RC involving matrices of 25 (5x5) and 36 (6x6) InGaAsP/InP microrings have been studied in [90] and [91] as well.

Spiking nonlinearity

Transient state dynamics can eventually result in self-pulsation, a dynamical regime emerging from the competition between free carrier and thermal nonlinearities in high Q-factor microring resonator [84][65][85]. The microring response to a CW optical signal is here characterized by a series of spikes, both at the through and drop ports. This is shown in in Fig. 3.6(a)(b), for a microring radius of $r = 6.75\mu m$, with $FWHM \sim 138pm$ and $Q \sim 11 \times 10^3$, whose resonance is shown in Fig. 3.6(d) as a function of the starting wave-

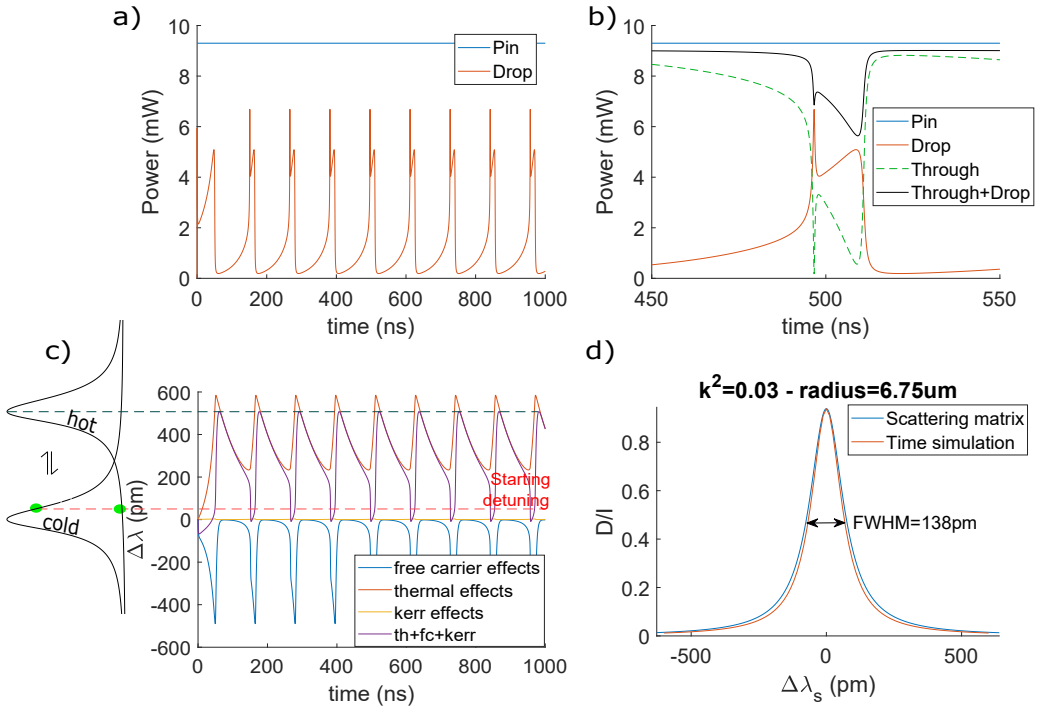

Figure 3.6: Example of simulated self pulsation dynamics achieved with a microring having radius $r = 6.75\mu m$, $FWHM = 138pm$, $Q = 11 \times 10^3$, and $k^2 = 0.03$, under CW optical injection at the input port, with $P_{in} = 9.3mW$. a) Drop temporal response and b) zoom, showing additionally the complementary through response. c) nonlinear contribution of the thermal, free carrier, and Kerr effects to the microring resonance shift, d) Linear drop transmission of the simulated microring.

length detuning $\Delta\lambda_s = \lambda_p - \lambda_0$.

To understand the mechanism behind self-pulsation, let's remember that the opposite changes in refractive index induced by thermal (Eq. 2.19) and free carrier (Eq. 2.23) nonlinearities can be observed as a shift of the resonance position over time (Eq. 3.16). Thus, the reader can refer to Fig. 3.6(c), for each thermal, free carrier, and Kerr nonlinear contribution $\Delta\lambda$ to the resonance shift, and for the effective shift given by their sum (purple line), while the spike emission in Fig. 3.6(a) is occurring. The simulation refers to a microring having radius $r = 6.75\mu m$, $FWHM = 138pm$, $Q = 11 \times 10^3$, $k^2 = 0.03$, starting detuning $\Delta\lambda_s = 50pm$ (red dashed line in Fig. 3.6(b)),

and to an optical power at the input port $P_{in} = 9.3mW$. When the carrier signal is switched on, the resonance is initially red-shifted due to thermal heating. But the more the resonance is red-shifted, and a lower optical power is coupled to the microring. The microring starts then to cool down, and the resonance blue shifts approaching λ_p. The input power within the microring starts to increase again, up to a point where a large free carrier population is generated, due to TPA nonlinear processes, and a strong blue-shift occurs. The free carrier effects is followed again by thermal heating that finally leads to a new red-shift of the resonance so that the loop is repeated. A spike is generated at each iteration. Note that each spike at the drop port is composed of two maxima, originating when the wavelength λ_p becomes perfectly resonant, firstly because of the free carrier blue-shift, and then because of thermal effects. Moreover, the second maximum is always lower than the first one, reflecting a larger population of free carriers at this time and thus larger losses by FCA. The reader can refer to Fig. 3.6(b) for a zoom of a spike, and for the complementary dynamic that is observed at the through port. The sum of the drop power and the through power reflects the losses by TPA.

Not all the starting wavelength detuning conditions lead to self-pulsing behavior. Self-pulsation is in principle more difficult to achieve at negative detuning, where a larger free carrier nonlinearity is initially needed to trigger the phenomena. If this is not the case, because for example the microring Q-factor is not sufficiently high, thermal effects dominate and the optical limiting phenomena are observed for negative detuning.

The frequency of self-pulsation depends on the time the microring takes to cool down, between consecutive spikes, before the input optical power can be coupled again to the microring triggering the free carrier action, that in turn reiterates the spiking process. Thus, the thermal timescale plays a significant role in determining the spiking frequency. Moreover, the higher the input optical power injected to the microring, the larger the resonance is red-shifted because of heating. As a result, the resonance cools down towards the pumping wavelength in a longer time period. The self-pulsing frequency change, as a function of the input optical power, is shown in Fig. 3.7, where the relationship has been investigated for some positive and small negative starting detuning. All curves are proper of a second-class neuron since the discontinuity occurs close to the self-pulsing threshold [92]. At the same time, some starting detuning conditions are more sensitive to the variation of the input power, as they lead to a wider range of self-pulsing frequencies. They

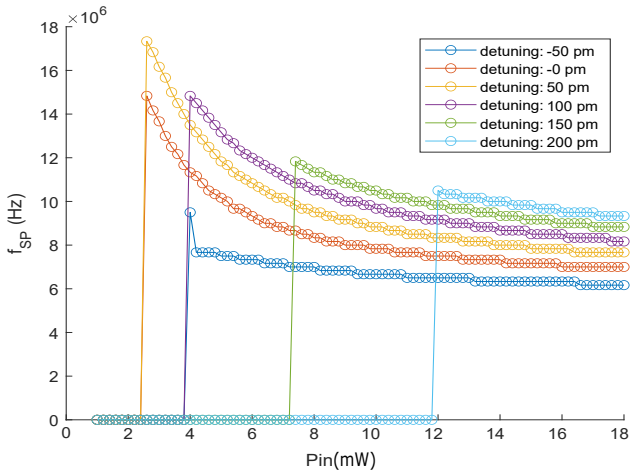

Figure 3.7: Self pulsing frequency (f_{sp}) as a function of the input power, for different starting detuning. The simulated microring has radius $r = 6.75\mu m$, $FWHM = 138pm$, $Q = 11 \times 10^3$, $k^2 = 0.03$.

catch in this way a peculiarity of class 1 biological neurons, whose spiking frequency depends on the strength of the applied current [93]. At more negative detuning conditions than the one plotted in Fig. 3.7, only one spike is generated if the input power is sufficiently high, right after switching on the input light. This spike is followed by oscillations that lead to a stationary solution, thus mimicking phasic spiking biological neurons [93].

Self-pulsation is an attractive feature for spiking neural networks. In [94], properties like the self-pulsing threshold, excitability, refractory time, and cascadability are studied in detail, by driving a microring on the onset of self-pulsation using a CW optical signal. As a result, the microring becomes more easily excitable when a second probe signal encoding a short input optical pulse is input in parallel: when the probe pulse carries enough energy, the microring emits a spike. Interestingly, the authors show experimentally that the spike emitted by a microring may serve as a perturbation for a second microring, still kept on the onset of self-pulsation. This demonstrates the cascadability of the self-pulsation between coupled microring resonators. The refractory period is another important property of a biological neuron, that inhibits the effects of any input signal within a defined time window that follows the spike emission. Self-pulsation also replicates the refractory period. Indeed, when the resonance is pushed far apart from the pump wavelength,

the microring is not sensible to any input perturbation involving the same λ_p. Note that when scaled to large SNNs, the refractory period also directly determines the operation speed of the whole network.

Chapter 4

Silicon microring resonators for time-delay reservoir computing

Microring resonators have been considered in the precedent section for introducing optical nonlinearity in photonic neuromorphic systems. In this chapter, a detailed numerical investigation of a delayed-silicon microring resonator for time delay RC is presented. The versatility of this passive device is demonstrated, by exploiting different operating regimes and solving computing tasks with diverse memory requirements. All the results are referred in [95].
The chapter is organized as follows. Firstly, the basic concepts of a time delay RC system are introduced in section 4.1. In section 4.2 these concepts are adapted to the specific case of a microring-based time delay RC. Here, the numerical modeling of the system, together with strategies to adapt it in a time-delay RC scheme, are described. Then, in section 4.3 the results obtained for different benchmark tasks are demonstrated. Finally, the effects of noise applied at the readout layer are investigated in section 4.4.
The results obtained in this chapter provide the basis to understand time-delay RC using microring resonators and are preparatory for the design and realization of the experimental investigations that will be described in chapters 5 and 6.

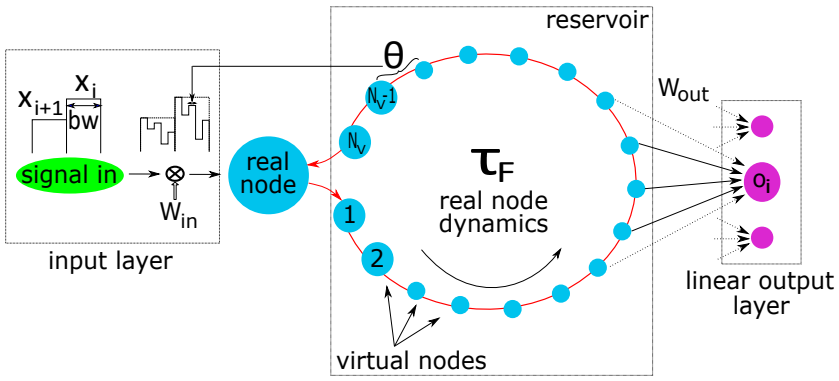

Figure 4.1: A time-delay RC scheme involves an input and an output layer, and a reservoir folded in the dynamics of only one node.

4.1 Time delay RC

4.1.1 Basic concepts

In 2011 Appeltant et al. [4] introduced an attractive strategy for physically implementing analog ANNs in a practical hardware setup, with a relatively high number of neurons. The scheme, which is sketched in Fig. 4.1, resembles the RC one (section 2.3). It consists of three parts: the input layer, the reservoir, and the output layer. The main idea consists of folding an ANN in the nonlinear dynamical response of a single node, named *real* node, and simply operating a linear regression on its different temporal responses. The nonlinear dynamics of the node is essential to map the information at the input layer into a higher dimensional space, given by the number of sampled temporal responses. These last emulate the reservoir nodes. Like in RC, only the output weights (projecting to the output layer) are trained during a task. Note that the network nodes are usually referred to as *virtual nodes*, to highlight that are time-multiplexed in the response of the real node and do not have any physical correspondence. Otherwise sad, the response of the reservoir is sampled multiple times as a function of time rather than as a function of space, like instead occurs in traditional RC.

The processing scheme is typically referred to as time-delay RC (or delay-based RC) because of the wide application using delay-based systems, since its introduction. This is also accounted for in Fig. 4.1, with a response signal which is fed back to the real node. When considering a continuous

time variable, delay-based systems can be described, in a simple form, by the following delay differential equation:

$$\dot{\mathbf{a}}(t) = f\left[t, \mathbf{a}(t), \mathbf{a}(t - \tau_F)\right], \tag{4.1}$$

where $\mathbf{a}(t)$ is the state of the real node at time t, $f(\cdot)$ is a nonlinear function, and τ_F is the feedback delay. Mathematically, a continuous time-delayed system is infinitely dimensional, as the initial conditions need to be specified over the interval $[-\tau_F, 0]$. Practically, their initial state can be approximated by a discrete signal sampled at least twice faster than the response time T of the system [96]. Assuming $T << \tau_F$, a high number of virtual nodes can be defined, and the real node is operated in a transient state.

Time-multiplexing is adopted at the input layer to serialize the information to be processed. Each input information (x_i) has to be stretched in time for a duration that will be indicated from now on as b_w (bit width). The duration b_w can be chosen differently according to the real node characteristic timescale T. When the input bit is multidimensional, i.e. $x_i = (x_{i,1}, x_{i,2}, \dots)$ its components can be either inserted in parallel to the system, when more input channels are available, or inserted along the same input line, by multiplexing in time the components of x_i. The scheme in Fig. 4.1 represents the case of unidimensional input bits x_i.

An additional mask procedure is commonly applied on top of the stretched input information for introducing transient dynamics in the response of the real node. The mask refers to an additional modulation that is periodically applied within each bit duration b_w, whose timescale depends on the characteristic time T of the real node. The generated transfer dynamics serve to increase the richness of the dynamics and, hence, to increase the range of values that virtual nodes can assume.

4.1.2 Spatiotemporal topology analogy

Time delay RC uses input time-multiplexing to serialize the information to be processed and combines it with a masking procedure to enrich the dynamical response of the real node, by introducing coupling between virtual nodes. The network topology and the coupling strength between the virtual nodes emerge from the relation of the real node characteristic timescale T and the mask timescale θ, and from the feedback recursivity. Based on these elements, different coupling strategies can be realized.

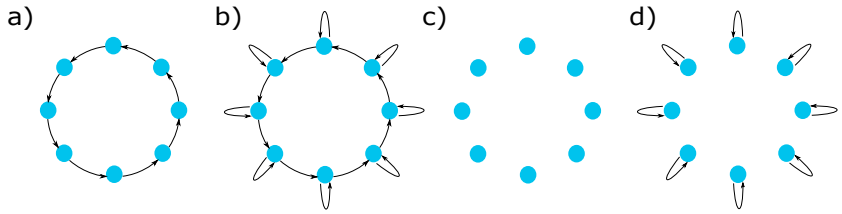

Figure 4.2: Coupling possibilities between virtual nodes in time-delay RC.

A first possibility is encountered when $\theta < T$. In this case, the mask allows to keep the system in a transient state: the system response can not reach a steady state before the next mask value is input. As a result, the virtual nodes, whose sampling time is θ as well, are coupled in a ring topology (Fig. 4.2(a)). The presence of a feedback line allows for additional couplings between the virtual nodes. If the input bit is synchronous with the delayed bit, i.e. $b_w = \tau_F$, with τ_F the feedback delay, each virtual node is also coupled with its state at the previous clock cycle (Fig. 4.2(b)). Otherwise, when $b_w \neq \tau_F$ neighboring virtual nodes, already coupled by the transient dynamics of the real node, become also connected via the feedback recursivity.

A second possibility is that $\theta >> T$. This is the case of a fast nonlinear response with respect to the input mask modulation period. Practically, this situation is encountered when the electronic instruments available to generate the mask signal are much slower than the real node nonlinearity timescale [97]. Here, the system reaches a steady state for each value of the input mask. If there is no feedback, each virtual node state is totally independent, and therefore there isn't any formation of a network topology. (Fig. 4.2(c)). On the other hand, when the feedback is applied in a synchronous regime ($b_w = \tau_F$), each node becomes coupled with its state at the previous cycle (Fig. 4.2(d)). Anyway, the coupling between different virtual nodes is still lacking. The reservoir states correspond indeed to the response of uncoupled dynamical systems, which slightly differ because of the different mask values. To couple the virtual nodes, it is essential to use an asynchronized feedback regime, with $\tau_F - b_w = k\theta$ [97]. Thanks to this method, each virtual node becomes coupled with the k^{th} previous virtual node. In the case $k = 1$ (the most common value used), a ring topology again emerges (Fig. 4.2(a)). Nevertheless, the ring topology essentially differs from the one obtained with a transient dynamics ($\theta < T$), since the coupling now involves the states of

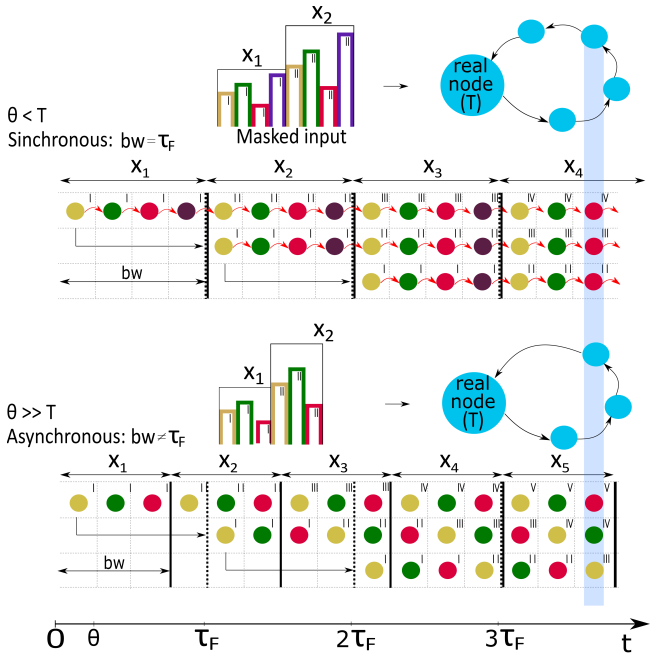

Figure 4.3: Two coupling possibilities between the virtual nodes, relying on the feedback signal. In a synchronous regime $bw = \tau_F$, with $\theta < T$, virtual nodes are coupled through the feedback to themselves at previous τ_F cycles (vertical light blue line), and between them by the real note inertia (red arrows). In the asynchronous regime $\tau_F - bw = \theta$, with $\theta >> T$, virtual nodes are only coupled to the state of their previous virtual node at previous τ_F cycles (vertical light blue line). A ring topology emerges in both cases, via the real node inertia in the first one and via the feedback recursivity in the second one.

virtual nodes belonging to different clock cycles. To grasp this last fact, an illustrative example is reported in Fig. 4.3, for the two cases $\theta < T$ and $\theta >> T$, the former in the synchronous regime and the second in the asynchronous regime. The mask values are here reported with different colors, with a Roman number indicating which input x_i they refer to. In the example, the synchronous regime refers to $\tau_F = b_w = 4\theta$, while the asynchronous regime refers to $\tau_F = 4\theta$ and $b_w = 3\theta$. In both examples, different rows highlight the information that is contributing to the system response. For $0 < t < \tau_F$ the node response only depends on the actual input information. For $\tau_f < t < 2\tau_F$, the real node response accounts for two contributes, the

actual input information and the previous one, which is delayed by one clock cycle τ_F, and input again to the real node. The same rule applies to the case $2\tau_f < t < 3\tau_F$. The two illustrated examples show how the information is coupled in the two respective regimes.

4.1.3 Delay-based implementations

Delay-based systems are particularly suitable for time-multiplexing reservoirs, and for this reason, they have been widely investigated since the introduction of this processing scheme. From the first electronics implementation in 2011 [4], many optoelectronics and all-optical implementations followed. A list of many delayed systems reported in the literature is presented in Table 4.1. The reader can refer to [96] and [98], for a detailed review regarding the implementations reported.

Type	nonlinear node	asynch.	Feedback	N_V	τ_F	Ref.
OE	Mach Zehnder	k=1	SMF	50-200	$8.5\mu s$	[97]
OE	Mach Zehnder	k=0	SMF	400	$20.87\mu s$	[99]
AO	SOA	k=1	SMF	50	$7.9\mu s$	[100]
AO	SRL	k=0	SMF	200	$4ns$	[101]
AO	SL	k=0	SMF	388	$77ns$	[102]
AO	SESAM	k=1	EDFA	50	$8\mu s$	[103]
AO	fiber cavity	k=1	SMF	50-500	$1.13\mu s$	[104]
AO	VCSEL	k=0	OC	325	$65ns$	[105]
AO	DBF	k=0	WG + EM	6	$254ps$	[106]
AO	DBR	k=0	InP WG	23	$1.170ns$	[107]

Table 4.1: Time-delay RC implementations. SMF=Single Mode Fiber; OE=Optoelectronics; SOA=Semiconductor Optical Amplifier; SRL=Semiconductor Ring Laser; SL=Semiconductor laser; SESAM=Semiconductor Saturable Absorber Mirror; VCSEL: Vertical Cavity Surface Emitting Laser; OC=Optical Circulator; AO=All Optical; DBF=Distributed Feedback semiconductor laser; DBR=Distributed Bragg laser; WG=waveguide; EM=External Mirror

In the first optoelectronics implementations, the real nonlinear node is implemented with a voltage-driven Mach Zehnder Interferometer (MZI), supplied by a CW optical signal, introducing a sine nonlinearity on it. The op-

tical signal propagates then through the feedback line, at the end of which is photodetected, and combined with the electronics input to the MZI, closing in this way the loop. The MZI nonlinearity applies essentially instantaneously on the propagating light, as it emerges from the interference condition at the output of the MZI. Hence, it is hard to induce nonlinear transient dynamics by using an electronic mask signal to the MZI. Thus, an experiment exploiting an asynchronous regime ($\tau_F - b_w = k\theta$, with k=1) has been proposed [97]. Alternatively, another group introduced a low pass filter after the detection step to provide the system with a characteristic timescale T. In this case, an asynchronous regime (k=0) was studied for time-delay RC [99]. The same system scheme was also studied in [108], with an optical switch inserted prior to the photodetection. The authors were able to compare the performance of the system on benchmark tasks with feedback and without feedback, eventually compensating for the lack of feedback recurrency by introducing artificial memory in the input information.

Other nonlinear nodes have been investigated, including Semiconductor Optical Amplifier (SOA) [100], Semiconductor Ring Laser (SRL) [100], Semiconductor Laser (SL) [102], Semiconductor Saturable Absorber Mirror (SESAM) [103], Vertical Cavity Surface Emitting Lasers (VCSELs) [105] and a passive fiber cavity [104]. When using a passive fiber cavity, the nonlinearity exploited is directly the output detection square nonlinearity, which measures the network response, before computing offline the output layer. All these implementations consider delay lines based on optical fibers, and are supported by an offline training that requires the exchange of data with a conventional computer, with the output weights typically optimized by a linear ridge regression algorithm. A proposal of a stand-alone time delay RC was also proposed in [109]. The application relies on the same optoelectronics implementation based on a delayed-Mach Zehnder interferometer already mentioned, which communicates now with a field-programmable gate array (FPGA). This last provides both the electrical input information to the MZI and acquires the output photodiode signal. The FPGA is also programmed to perform an online gradient descent for the optimization of the output weights.

Multiple delay lines have been also considered. An example is given in [110], where the authors model the possibility to unfold a multilayer deep neural network, in the dynamics of a single nonlinear node subject to multiple delay lines. For example, fully connected layers with N nodes are emulated using $N-1$ delay lines, each one equipped with a modulator, that emulate

the weights between the layers.

Other implementations are inspiring because performing multiple tasks at the same time. An example concerns the case of a SRL cavity [101]. The system is here suitable for computing two tasks in parallel, exploiting the clockwise and anticlockwise propagation directions of light in the SRL.

The integration of time-delay RC systems has been also explored, with integrated Distributed Feedback semiconductor lasers (DBF) [106] and Distributed Bragg Reflector (DBR) lasers [107] as nonlinear nodes. However, integrated implementations are in general limited by the length of the feedback loop which in turn limits the available number of virtual nodes ($N_V \approx \tau_F/\theta$). Offline techniques can help to compensate for the low number of virtual nodes, for instance by acquiring more virtual nodes per mask node, or by sampling the virtual nodes over multiple bw time intervals, but at the cost of a slower computational speed.

This argument highlights the importance of the choice of the mask. A piecewise function constant over intervals of length θ is the most common choice, despite analog masks, such as sinusoidal signals, have been used as well [111].

Delay-based systems in time delay RC have been tested on tasks including time series prediction, delayed-boolean tasks, spoken digit recognition, and nonlinear channel equalization. This last task is particularly investigated exploiting the nonlinear dynamics and bandwidth of SLs with a time-delay feedback loop, as reported in [112][113][114]. Here, bandwidth enhancement mechanisms are possible by optimizing the optical signal injection from the master laser. In this way, it is possible to shorten the virtual nodes sampling time down to 12 ps ($\tau_F = 8$ ns clock time), for faster signal recovery speed [115] that eventually matches real-time applications.

In the next part of the chapter it will be presented the numerical investigation of a novel real node for an all-optical time delay RC based on an integrated silicon microring resonator with external optical feedback

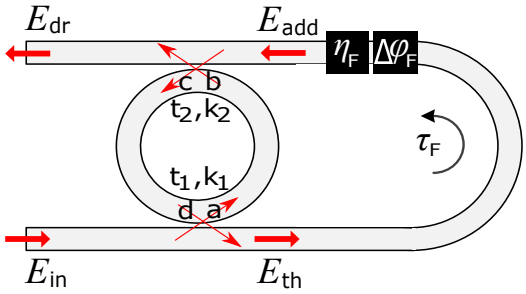

Figure 4.4: Microring in an add-drop filter configuration with external optical feedback. E_{in}, E_{th}, E_{ad} and E_{dr} represent the electric field amplitudes in the correspondent ports, while η_F and $\Delta\phi_F$ are parameters related to the feedback intensity (strength) and phase, respectively.

4.2 Microring resonator with external optical feedback as real node

The delayed system of the proposed time-delay RC implementation is a silicon microring resonator coupled to an external optical feedback [95]. The microring is designed in add-drop filter configuration, and the feedback line is realized by connecting the through and add ports, as shown in Fig. 4.4. This section describes the numerical model of the system, as well as how to adapt it in a time-delay RC implementation.

4.2.1 Real node modelling

Linear frequency response

A scattering matrix approach can be applied to the system, similarly to section 3.1.1, to obtain the linear frequency response. The equations that are reported below take into account the feedback contribution.

$$\begin{pmatrix} E_{dr} \\ c \end{pmatrix} = \begin{pmatrix} ik_2 & t_2 \\ t_2 & iK_2 \end{pmatrix} \begin{pmatrix} b \\ E_{ad} \end{pmatrix} \tag{4.2}$$

$$b = \sqrt{a_{rt}} e^{-i\phi_{rt}/2} a \tag{4.3}$$

$$E_{ad} = \sqrt{\eta_F} e^{-i\phi_F} E_{th} \tag{4.4}$$

$$\begin{pmatrix} a \\ E_{th} \end{pmatrix} = \begin{pmatrix} ik_1 & t_1 \\ t_1 & iK_1 \end{pmatrix} \begin{pmatrix} E_{in} \\ d \end{pmatrix} \tag{4.5}$$

Along the feedback line, two additional parameters η_F and $\Delta\phi_F$ are introduced to control the strength and the phase of the feedback signal, respectively. The parameter $\eta_F \in [0, 1]$, with zero value assigned to the complete attenuation of the signal and one assigned to the complete transmission of the feedback signal. The parameter $\Delta\phi_F \in [0, 2\pi]$ is a tunable phase quantity that is added to the phase of the feedback signal, so that $\phi_F = \beta_F L_F + \Delta\phi_F$, with L_F the length of the feedback line and β_F the correspondent propagation constant. These parameters, as will be shown in the following, will be optimized when testing the system on computational tasks.

The drop port transmission $D = |E_{dr}/E_{in}|^2$ can be derived in the following way. First, E_{dr} is expressed as a function of E_{in} and d by concatenating Eqs. 4.2-4.5. Then, the expression of $d = \sqrt{a_{rt}} e^{-i\phi_{rt}/2} c$ is updated with the expression of the field c given by Eq. 4.2, to become a function of E_{in} as well. Hence, the optical drop transmission can be calculated as:

$$D = \left|\frac{E_{dr}}{E_{in}}\right|^2 = \left|\frac{t_1 t_2 \sqrt{\eta_F} e^{-i\phi_F} - k_1 k_2 \sqrt{a_{rt}} e^{-i\phi rt/2} - \sqrt{\eta_F} a_{rt} e^{-i(\phi_F + \phi_{rt})}}{1 - \left[t_1 t_2 a_{rt} e^{-i\phi_{rt}} - k_1 k_2 \sqrt{\eta_F}\sqrt{a_{rt}} e^{-i(\phi_F + \phi_{rt}/2)}\right]}\right|^2, \tag{4.6}$$

that simplifies to Eq. 3.8 when the feedback signal is fully attenuated ($\eta_F = 0$).

Nonlinear time response

The temporal response of the microring coupled to an external feedback follows the one of a single microring in add-drop filter configuration, already reported in Eq.3.15, but considering an additional input field also at the add port (E_{ad}). The latter is included in the equation of the internal energy amplitude as an additional driving force. For symmetry reasons, the equation describing the dynamics of the internal field amplitude $\tilde{U}(t)$ is written as:

$$\frac{d\tilde{U}(t)}{dt} = \left[-i\left(w_p - \left(w_0 - \delta w_{nl}^R(t)\right)\right) - \left(\frac{1}{\tau_{tot}} + \Delta w_{nl}^I(t)\right)\right]\tilde{U}(t) + i\sqrt{\frac{2}{\tau_k}}(E_{in} + E_{ad}). \tag{4.7}$$

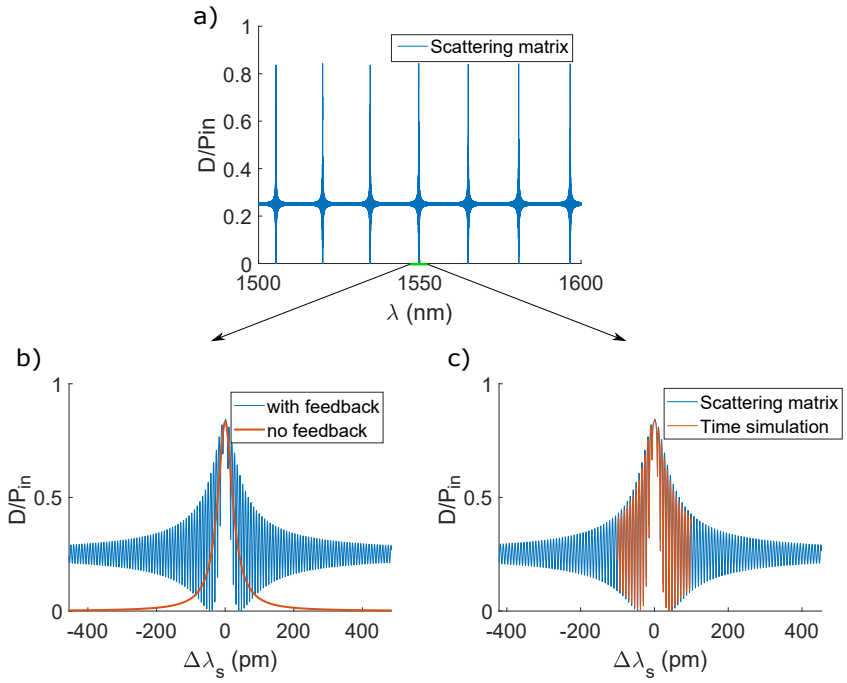

Figure 4.5: Drop transmission of a microring resonator, with a feedback connecting the through and add ports. The microring has a radius $r = 6.75\mu m$, symmetric coupling coefficient $k_1 = k_2 = k$, with $k^2 = 0.01$, a feedback delay $\tau_F = 1ns$, a feedback strength $\eta_F = 0.25$, and no control over the feedback phase ($\Delta\phi_F = 0$). a) Wide drop spectrum. b) Comparison of the spectra in presence and absence of the feedback. c) Comparison of the spectra obtained from a scattering matrix approximation (Eq. 4.6) and from a numerical simulation of the temporal solution (Eq. 4.7).

From Eq. 4.7, the drop response is calculated as:

$$E_{dr} = t_2 E_{ad} + ik_2 \frac{\tilde{U}(t)}{\sqrt{p/v_g}}, \tag{4.8}$$

with $E_{ad}(t) = \sqrt{\eta_F} e^{-i\phi_F} E_{th}(t - \tau_F)$, being τ_F the feedback delay, p the microring perimeter and v_g the group velocity.

The use of Eq. 4.7 is validated by simulating the drop linear response for many input wavelength detuning values, thus extracting the correspondent steady states normalized with respect to P_{in}, and then comparing those with the linear transmission expected from Eq. 4.6. Note that the time needed to

reach a steady state depends now on the delay provided by the feedback line. Thus, the system needs to be simulated for a time longer than τ_F to reach its steady state. The results that follow refer to the drop transmission obtained for a microring radius $r = 6.75\mu m$, a feedback delay time $\tau_F = 1ns$, and a simulation time of $2\mu s$ ($>> \tau_F$). Fig.4.5(a) shows the drop transmission in presence of feedback for a wide region of input wavelengths, while Fig. 4.5(b)(c) is focused around one of the microring resonances. In particular, Fig. 4.5(b) compares the transmission of the microring when including or not the feedback. When the feedback is not present, the drop signal reduces to almost zero for large detuning $\Delta\lambda_s$, as light propagates to the through channel. The through signal, if the feedback line is connected, is reinjected back to the add port and then to the drop port. As a result, the drop transmission for large detuning $\Delta\lambda_s$, is non-zero and depends on the feedback strength ($\eta_F = 0.25$ in the example). If $\Delta\lambda_s$ is instead partially in resonance, larger interference fringes at the drop port appear. Fig. 4.5(c) shows the agreement between simulation and scattering matrix spectrum, validating the use of Eq. 4.7.

4.2.2 Feedback effects on nonlinearity

Both the amplitude and phase of the optical feedback signal are quantities that affect the optical field within the microring, and hence its nonlinear dynamics. Indeed, the delayed feedback signal can be coupled again to the microring, as represented in the inset of Fig. 4.6, and then interfere with the internal optical field.

Three main situations can occur, assuming the microring in a linear regime and a CW input signal. (i) The optical input signal is perfectly resonant ($\Delta\lambda_s = 0$), and almost no signal propagates through the feedback line. (ii) The optical input signal is out of resonance ($|\Delta\lambda_s| >> FWHM$), so it bypasses the microring, reaching the drop port with a delay τ_F. (iii) The input optical signal is partially resonant ($\Delta\lambda_s \sim FWHM$) so that it is both coupled to the microring and to the feedback line. Similarly, at the add port, the optical feedback signal is partially coupled back to the microring and partially transmitted to the drop port. Thus, the feedback signal interferes with the microring's internal electric field. As a consequence, the tuning of the feedback phase $\Delta\phi_F$ and strength η_F, allows controlling the optical power within the microring by changing the interference condition. The example reported in Fig. 4.6 shows that the tuning of the parameter $\Delta\phi_F$ is

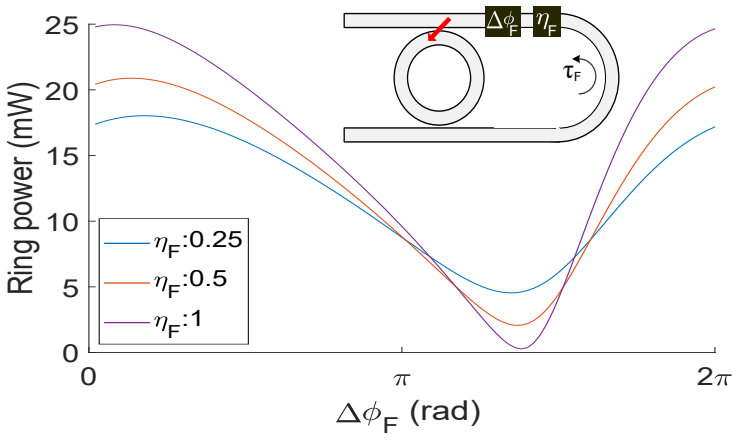

Figure 4.6: The delayed signal is coupled back to the microring. The feedback phase $\Delta\phi_F$ set the interference condition within the microring, while its strength η_F determines its contribution.

the means to select the interference condition within the microring, lowering the internal optical power with destructive interference, or enhancing it with constructive interference. As a result, $\Delta\phi_F$ and η_F represent additional tools for controlling the microring internal power, and thus the microring nonlinearity.

4.2.3 Processing scheme

The scheme of the proposed time-delay RC is sketched in Fig. 4.7. It follows the typical scheme, consisting of an input layer, the real node, and an output layer. The real node here is a silicon microring resonator coupled to an external optical feedback, connecting the through to the add ports with a delay τ_F. The microring receives an optical signal at the input port. Its response is encoded in the detected drop signal, where the virtual reservoir is folded.

The input layer properly encodes the information in the optical domain, before being input to the microring. This information can be either continuous in time or with discrete values, as discussed in [4]. In both cases, it is normalized in the range [0,1]. Then, sampled values x_i from the sequence are extracted and translated in bits of proportional amplitude and duration b_w. The resulting signal modulates the optical carrier emitted by a CW laser,

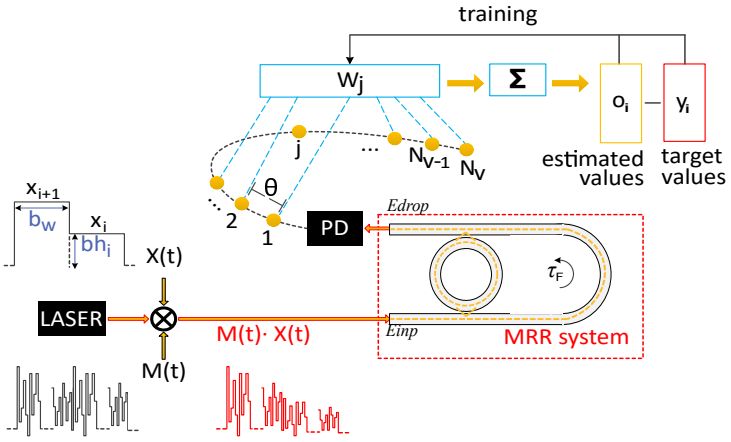

Figure 4.7: Schematic of time delay RC with an MRR subject to optical feedback. The encoded information $X(t)$ is masked with a sequence $M(t)$ and modulates the optical power from the laser (LAS) emission. At the drop port, the photodetected (PD) signal provides the time-multiplexed output states of the reservoir, which are weighted and linearly combined to compute the predicted value o_i. The weight optimization is performed via a linear classifier, with supervised learning over the expected values y_i data set.

with the amplitude of each optical bit then becoming $b_{h,i} = x_i P_{max}$, with P_{max} the power of the optical signal before the modulation. Each optical bit is additionally masked with a set of random values $M(t)$ taken from a uniform distribution, with $M(t) = M(t+b_w)$. Then, the optical signal enters the input port of the microring and propagates along the structure. Different paths (orange dotted line in Fig. 4.7) are allowed, depending on the detuning between the pump wavelength λ_p and the resonance wavelength $\lambda_0(t)$, as discussed in section 4.2. The optical response is detected at the drop port, from which the virtual nodes are sampled and the output estimation of the network o_i, relative to the input bit x_i, is computed. The nonlinear transformation of the input signal occurs at both the microring and the photodetection stages. Additionally, each uni-dimensional input information is eventually projected through this physical system into a higher dimensional space defined by the number of virtual nodes $N_v = \frac{b_w}{\theta}$, with θ referring to their temporal distance separation.
By defining $N_{j,i}$ as the j^{th} virtual node response associated with the i^{th} processed element x_i, the output of the network o_i is estimated as a linear

weighted sum, like in traditional RC. In formulas:

$$N_{j,i} =\mid E_{drop}(b_w(i-1)+\theta j) \mid^2, \tag{4.9}$$

$$o_i = \sum_{j=1}^{N_v} W_j N_{j,i}, \tag{4.10}$$

where W_j is the N_v-dimensional vector of the readout weights. The latter also represent the only parameters that are optimized during the training phase of a task, based on a chosen cost function to minimize. A cost function commonly used is the Normalized Mean Square Error (NMSE), defined as:

$$NMSE = \frac{\sum_i (o_i - y_i)^2}{N_d \sigma_y^2}, \tag{4.11}$$

where the sum includes all the elements of the test dataset N_d, σ_y^2 is the standard deviation squared of the test dataset, and y_i is the target output related to x_i. The lower the $NMSE$ in Eq. 4.11 is, the better the system predicts the expected output series. This metric will be applied later on in the numerical work to quantify the performance of the trained network on benchmark tasks. Finally, a linear ridge regression algorithm [116] is applied to find the optimal readout weights.

4.2.4 Parametrization of the system and dynamics

This section describes the choice of the optical bit duration b_w, mask function and timescale θ, and feedback delay time τ_F. As already mentioned, the dynamics of a microring resonator is influenced by the photon lifetime τ_{ph}, the free carrier lifetime τ_{FC} and the thermal lifetime τ_{TH}. Note that while τ_{ph} refers to the microring internal field discharge time, which introduces linear dynamics to the microring, τ_{FC} and τ_{TH} refer to the nonlinear dynamics triggered by TPA processes within the microring.

A first problem to face is how to manage these timescales for time-delay RC, considering that generally $\tau_{ph} < \tau_{FC} < \tau_{TH}$. In particular, free carrier nonlinear effects (here $\tau_{FC} \approx 3$ns) are typically two orders of magnitude faster than thermal effects (here $\tau_{TH} \approx 83$ ns). Since the nonlinearity timescale is directly linked to the computing speed, the choice is to process the input

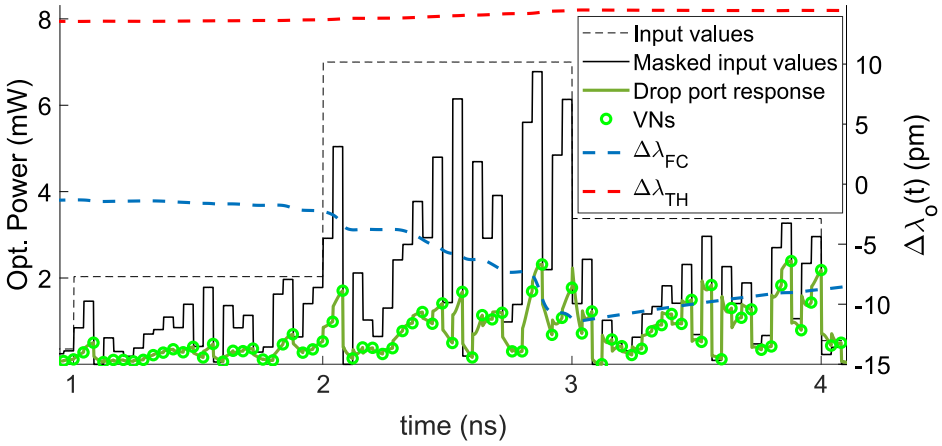

Figure 4.8: Dynamical response of a microring to a fast modulating signal at the input port, based on the input layer concept of a time delay RC scheme. Black dashed line and black continuous line represent the input information before and after the masking, respectively. The response of the microring is obtained at the drop port (green line), while the green circles indicate the sampled virtual node states of the reservoir, separated in time by $\theta = 40$ ps (one virtual node per mask node). The blue and red dashed lines show the contribution of the free carrier ($\Delta\lambda_{FC}$) and the thermal ($\Delta\lambda_{TH}$) nonlinear effects, respectively, to the wavelength resonance shift (measured in the right y-axis).

information by exploiting the free carriers of the microring, which have a faster response time. Thus, a microring exhibiting free carrier effects has to be designed. The simulated microring has a radius $r = 6.75\mu m$, a $FWHM = 48pm$, a quality factor of $Q = 3.19 \times 10^4$, and can exhibit self-pulsation dynamics, as discussed in section 3.3.2. Each optical pulse is stretched in time for a duration similar to the free carrier lifetime. Here, the choice of $b_w = 1ns$ is made, a time sufficiently large to allow the excitation of extra free carriers within the microring waveguide. The reader can refer to Fig. 4.5 for the linear drop transmission of the microring.

The fastest characteristic response time of the microring - associated with the photon lifetime (here, $\tau_{ph} = 52ps$) - is also exploited to keep the system's operation in a transient state, via the mask applied at the input layer. The letter corresponds to a piece-wise function, constant over intervals $\theta = 40ps$, periodically applied on each optical pulse. Thus, the microring does not

completely release the internal field, when the next mask sample is applied. This allows the coupling of neighboring virtual nodes through this kind of inertia. As a consequence, an asynchronous regime is not necessary to couple the virtual nodes. Therefore, a synchronous regime is applied by setting $bw = \tau_F$.

Note that the same time interval θ is used both for the mask values duration as well as for the virtual node sampling time, thus creating a one-by-one correspondence between mask node and virtual node. In this way, a total number of $N_v = bw/\theta = 25$ virtual nodes are defined within a duration of b_w. This number of virtual nodes allows computations at GHz rates and is also compatible with delays provided by integrated silicon feedback waveguides. Figure 4.8 combines all the concepts just discussed, showing both the masked input and the temporal response of the microring with an external feedback loop. The nonlinear contributions to the resonance shift induced by free carriers ($\Delta\lambda_{FC}$) and thermal ($\Delta\lambda_{TH}$) effects are also shown (blue and red dashed line, respectively). As can be seen, the microring resonance position blue shifts when the largest pulse is input to the microring, thus manifesting the presence of excited free carriers. On the contrary, thermal effects are much slower and do not participate in the nonlinear transformation of the input information. Still, they add a positive bias to the resonance position, and become important when self-pulsing dynamics are activated for sufficiently high optical power.

4.3 Task results

Under the above conditions, the microring-based network is tested on three different benchmark computational tasks, with different requirements for signal processing. The NARMA 10 belongs to the category of nonlinear system identification tasks and requires explicitly 10 memory steps to be solved [117]. The Mackey-Glass [118] and the Santa Fe [119] are benchmark one-step-ahead chaotic time-series prediction tasks where the system has to predict a future value x_{i+1} of the input series while processing x_i.

Since the performance of every task relies on a characteristic trade-off between a nonlinear transformation of the input information and the system's linear memory [120], an estimation of these two quantities is addressed below. The nonlinear transformation is evaluated indirectly, via the time dependence of the standard deviation of the wavelength resonance shift $\sigma(\lambda_o(t))$. A high

standard deviation indicates higher resonance shifts occurred along the task, and thus a higher microring nonlinearity. The system's linear memory is evaluated by the linear memory capacity (MC) task [121]. Traditionally, MC is calculated by using an input random sequence of bits, with values taken from a uniform distribution [122]. However, there is an inconsistency in this method when one tries to evaluate the response of a nonlinear system to an input with specific spectral properties. The linear memory capacity of the system can be different when entering into the system either a random sequence or a sequence with correlated temporal profiles. This stands also here, where different input series may lead the microring operation under different nonlinear dynamics, even for the same operating parameters. For this reason, here, the MC task is solved for the actual time series of the benchmark tasks that will be evaluated. The system is trained to remember the l^{th} previous input element of the used series, by exploiting the information that is still present in the system. In fact, MC is computed as:

$$MC = \sum_{l=0}^{l_{max}} m(l), \quad \text{with} \quad m(l) = \frac{cov^2(o(n), x(n-l))}{\sigma_o^2 \sigma_x^2}, \tag{4.12}$$

where $m(l)$ measures the normalized linear correlation between the predicted ($o(n)$) and delayed ($x(n-l)$) input series, $cov^2()$ indicates the covariance between two vectors and σ^2 the variance. When this correlation is very small ($m(l) << 1$), the system is unable to preserve any information of l past input. On the contrary, when $m(l) \approx 1$, the system remembers the exact value. For the MC computation, it is also considered the case of $l = 0$ which refers to the capability of the system to retrieve the actual input.

The performance of the microring with external optical feedback RC is investigated on the selected tasks by tuning the input power P_{max}, the starting wavelength detuning between the signal wavelength and the cold resonance wavelength $\Delta\lambda_s$, the feedback strength η_F and the feedback phase $\Delta\phi_F$. All these parameters affect the dynamics of the real node, by changing the nonlinear operation of the microring and the contribution of the feedback signal. Thus, the nonlinearity and the memory of the system can be probed. The following ranges of values are selected for these parameters: $P_{max} \in [1$ mW, 8 mW], with step of 1 mW, including also the value 0.1 mW where the microring operates in a linear regime; $\Delta\lambda_s \in [-50$ pm, 50 pm], with step of 10 pm, so that all the microring resonance (having $FWHM = 48$pm) is covered; $\eta_F \in [0, 1]$; $\Delta\phi_F \in [0, 2\pi]$.

To train the network, the discrete input series of each task are separated in the following way: 1000 input data values are used to drive the system in a working regime and eliminate any oscillatory operation due to the inclusion of the input. Then, 2000 input data values serve the training phase, and the next 1000 data the testing phase, where the system performance are tested on previously unseen entries. The latter are evaluated according to the NMSE, as defined in Eq. 4.11. The same random mask $M(t)$ is used in all the simulations that involve the same number of virtual nodes, which is fixed to 25 unless differently specified. The ridge regression parameter of the RC's output layer linear classifier is set to 10^{-4}.

4.3.1 Narma 10 benchmark task

In the Narma 10 task, the system is trained to predict the response of a discrete-time tenth-order nonlinear auto-regressive moving average (NARMA) system [117], described by the following equation:

$$r_{i+1} = 0.3r_i + 0.05r_i(\sum_{j=0}^{9} r_{i-j}) + 1.5x_{i-9}x_i + 0.1, \tag{4.13}$$

where x_i represents the i^{th} element of the input series uniformly distributed in the range [0, 0.5] and r_{i+1} is the correspondent expected target, labeled by y_i in the scheme of Fig. 4.7. Eq. 4.13 suggests that at least 10 values (the current one and 9 in the past) need to be considered to predict the next value.

In Fig. 4.9(a) the NMSE performance of the microring system is shown, versus the feedback parameters η_F and $\Delta\phi_F$. In parallel, the linear MC of the microring-based system is also shown for this task (Fig. 4.9(b)). While for the performance optimization, it is used the 4-dimensional parameter manifold (P_{max}, $\Delta\lambda_s$, η_F, $\Delta\phi_F$), the results are provided in two dimensions while fixing the rest of the parameters of the complete manifold. The parameter space for which the lowest values of NMSE are observed is where the MC approaches its maximum value. This in turn is achieved when the microring is operated in a linear regime. The results in Fig. 4.9, hence, are obtained for input power $P_{max} = 0.1mW$. Still, the network exhibits a nonlinearity, given by the photodetection.

First, I consider the absence of the external optical feedback in the system ($\eta_F = 0$). In this case, the MC value is only around 2 (Fig. 4.9(b)), and the

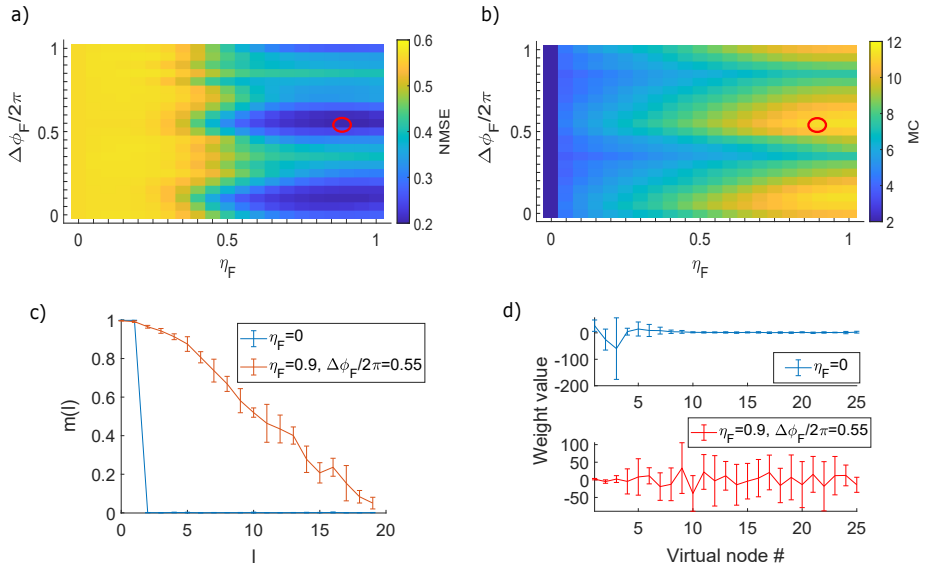

Figure 4.9: Performance of the neural network to the Narma 10 benchmark task. (a) NMSE and (b) MC, versus optical feedback strength η_F and phase $\Delta\phi_F$. The red circle denotes the conditions with the lowest NMSE. (c) Memory function $m(l)$, for the cases without feedback (blue line) and with feedback conditions that result in the lowest NMSE (red line). (d) Readout weights for a task to remember the previous input value x_{i-1}, for the cases without feedback (blue line) and with feedback conditions that result in the lowest NMSE (red line). MC is computed using $l_{max} = 19$. The initial wavelength shift is $\Delta\lambda_s = -10pm$ and the microring is operating in the linear regime, with $b_w = 1$ ns.

system is only able to preserve information regarding the actual bit x_i and the previous one x_{i-1}, as shown by the blue line in Fig. 4.9(c). Information of the previous bit here emerges from the inertia between the last virtual nodes' responses to the input value x_{i-1} and the first virtual nodes' responses to the next input value x_i. This can be verified by a simple memory task that aims to give the past bit. In this case, the computed weight values of the RC linear classifier are shown in Fig. 4.9(d). When the classifier is trained to provide as an output x_{i-1}, by considering the response of the reservoir to the actual input x_i, only the response of the first virtual nodes is important for computation (Fig. 4.9(d), blue line). But when activating the feedback (Fig. 4.9(c), red line), all virtual nodes contribute to the task (Fig. 4.9(d),

red line). Thus, to obtain an extended linear memory, a strong feedback parameter η_F is required, under an appropriate phase condition at the add port of the microring. Here, MC values beyond 10 are exhibited for η_F approaching 1, while at the same time lowering the NMSE (Fig. 4.9(a)). In particular, the minimum error $NMSE_{min} = 0.204 \pm 0.026$ is found at $\eta_F = 0.9$, and $\Delta\phi_F/2\pi = 0.55$, for $P_{max} = 0.1mW$ and a starting wavelength detuning of $\Delta\lambda_s = -10pm$.

Note that the combination of a linear microring and an optical feedback line acts like an analog shift register. If the microring is in resonance, part of the feedback signal is coupled back to the microring. Thus, an input light pulse can propagate multiple times through the fiber loop, providing a linear optical memory to the system. A worse performance is expected when the microring operates in a nonlinear regime since it progressively distorts the information [120].

The lowest obtained NMSE value is even higher than the one expected from a linear shift register ($NMSE_{SR} = 0.16$) [4]. The reason for this is the small number of virtual nodes. But since in this task the microring dynamics is not bounded to its nonlinearity, a longer bit duration b_w can be considered while preserving the same dynamical response and the same virtual node time separation $\theta = 40$ps. For example, when $N_v = 200$ and $b_w = 8$ ns, a $NMSE = 0.010 \pm 0.009$ is obtained, an equivalent performance to [4]. This improvement, compared to the linear shift register performance, is attributed to the square-law nonlinearity of the photodetection, since both cases exploit the same linear memory. This operation is consistent with the one presented in [104], where a linear external cavity with an optical fiber loop was used and from which the output optical signal was photodetected. For comparison, the use of the microring in absence of optical feedback, results in $NMSE = 0.545 \pm 0.001$, for $P_{max} = 2$ mW and $\Delta\lambda_s = 20$ pm. When a higher input power is applied, the microring nonlinearities deteriorate the linear MC, and also the performance on this task.

In conclusion, the single microring operating in the linear regime and in absence of feedback can preserve the previous value of input information, through the inertia. For the Narma 10 task, which has longer memory requirements, the external cavity is the main contributor to the linear memory capacity of the computing system.

4.3.2 Mackey Glass benchmark task

The Mackey-Glass input series is obtained by integrating in time the following equation:

$$\frac{dx(t)}{dt} = \frac{\alpha x(t-\tau)}{1+x(t-\tau)^{\beta}} - \gamma x(t), \tag{4.14}$$

whose solutions may be periodic, aperiodic, and chaotic, depending on the values of the parameters. Eq. 4.14 was initially considered in [123] to describe physiological diseases in the human body and later, in recurrent neural networks [118], as a benchmark timeseries for prediction. In the last case, a weakly chaotic behavior is commonly used, by numerically solving Eq. (4.14) with an integration step of 0.1, and the following parameter values: $\alpha = 0.2$, $\beta = 10$, $\gamma = 0.1$, and $\tau = 17$. In this investigation, a downsampling by a factor 30 is also applied, similarly to [124].

The results of solving this tasks are shown in Fig. 4.10, for the parameters $P_{max} = 5$ mW and $\Delta\lambda_s = -30$ pm. The results are shown in the feedback parameter space via η_F and $\Delta\phi_F$, with the best configuration of the task obtained for $\eta_F = 0.85$ and $\Delta\phi_F/2\pi = 0.6$ (black circle). These conditions result in an $NMSE_{min} = 0.0053 \pm 0.0005$, lower than $NMSE_{SR} = 0.01$ obtained by a linear shift register. Differently from the Narma 10 task, the optimal computing conditions here exploit the microring nonlinearity, as indicated by the corresponding standard deviation value of the microring resonance shift $\Delta\lambda_o(t)$ (Fig. 4.10(b)). Nevertheless, the region with the largest resonance shifts in Fig. 4.10(b) is related to higher NMSE in Fig. 4.10(a). The external cavity, besides its contribution to the extended fading memory, is beneficial in the following sense: different interference conditions between the feedback signal and the internal field of the microring are achieved by tuning the η_F and $\Delta\phi_F$ values. Thus, the feedback conditions also determine the circulating internal optical power in the microring and eventually its nonlinearity. For example, when $\eta_F > 0.3$ and $0 < \Delta\phi_F/2\pi < 0.45$, a constructive interference is observed in the microring, leading to higher values of $\Delta\lambda_o(t)$, thus higher microring nonlinearity (Fig. 4.10(b)) and degraded performance (Fig. 4.10(a)). For this region of the feedback parameter space, the high nonlinearity is able to shift the microring resonance enough to uncouple it from the input signal wavelength. This is illustrated in Fig. 4.10(c, upper panel, red line) from the evolution of the resonance shift $\Delta\lambda_o(t)$ in time while executing the computation. For the conditions indicated with the red circle in Fig. 4.10(a), a series of bursting spikes in the wavelength

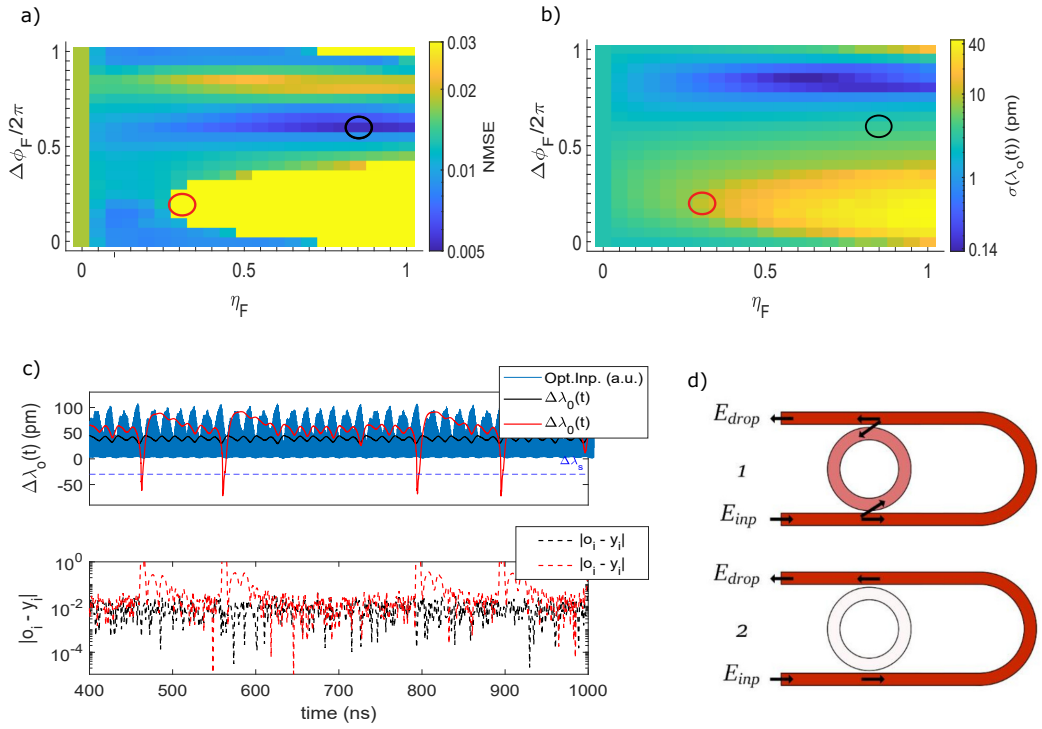

Figure 4.10: Performance of the Mackey-Glass benchmark task. (a) NMSE and (b) standard deviation of the resonance wavelength shift $\sigma(\lambda_0)$, versus optical feedback strength η_F and phase $\Delta\phi_F$ of the microring system. The black (red) circle denotes the conditions with the lowest (highest) NMSE. (c) Temporal evolution of the resonance shift and the bit error $|o_i - y_i|$ during the task for two feedback conditions: the black line corresponds to the lowest NMSE (black circle, (a)), and the red line corresponds to the highest NMSE (red circle, (a)). (d) Dynamical operation of the microring with optical feedback under self pulsations: light occasionally enters (path 1, upper) or bypasses (path 2, lower) the microring. The initial wavelength shift is $\Delta\lambda_s = -30pm$, the maximum launched optical power at the input is $P_{max} = 5$ mW and $b_w = 1$ ns.

resonance shift is observed, followed by a thermal warming and then a thermal cool-down. This dynamical behavior resembles self-pulsations, where the spiking emerges as a competition between the free carrier and thermal effects within the ring. Self-pulsing temporarily drives the resonance far away to the input wavelength (out of resonance condition). In these time intervals, the optical signal mainly propagates through the external cavity (Fig. 4.10(d, path 2)) and is not affected any more by the microring nonlinearity. At the same time, as the feedback signal passes the microring, it can not be coupled back and iterated further. In this way, the system loses the feedback memory. These conditions result in degraded performances, as it is indicated by the corresponding higher prediction error in Fig. 4.10(c, bottom panel, red line). As a comparison, the lowest NMSE configuration is also reported in Fig. 4.10(c, black line). In this case $\Delta\lambda_o(t)$ oscillates in phase with the input signal (Fig. 4.10(c, upper panel, blue line)) and with lower amplitudes. Note that in a comparison with the NARMA 10 task, the region with the worst NMSE performance (Fig. 4.10(a)) – where the microring transition to self-sustained oscillations occurs – differs from the one in Fig. 4.9(a), since in the latter the microring is operated in a linear regime. In a comparison with the linear shift register, neither the single microring in absence of the feedback ($NMSE = 0.015 \pm 0.002$) nor the linear microring in presence of feedback ($NMSE = 0.0095 \pm 0.0009$) provide improved performance.

4.3.3 Santa Fe benchmark task

The Santa Fe benchmark test is another one-step-ahead time series prediction task. In this task, the input series is the optical power emitted by a far-infrared laser that operates in a chaotic regime [119].

While processing this dataset with a linear shift register results in a value of $NMSE_{SR} = 0.2$, with the microring-based RC lower NMSE values can be obtained. This is shown in Fig. 4.11(a), where the best NMSE for each pair ($\Delta\lambda_s$, P_{max}) are plotted. The standard deviation of the resonance shift $\sigma(\lambda_0)$ (Fig. 4.11(b)) and the feedback strength η_F (Fig. 4.11(c)), related to the configurations with the lower error in Fig. 4.11(a) are also presented.
While all the displayed configurations achieve errors lower than $NMSE_{SR}$, a joint evaluation of the three figures allows interpreting why the best results are obtained with a given parameter set. At $P_{max} = 0.1mW$, the microring works in a linear regime, as also indicated by the small values of $\sigma(\lambda_0)$ at this power (Fig. 4.11(b)). In this condition, the feedback strength η_F is

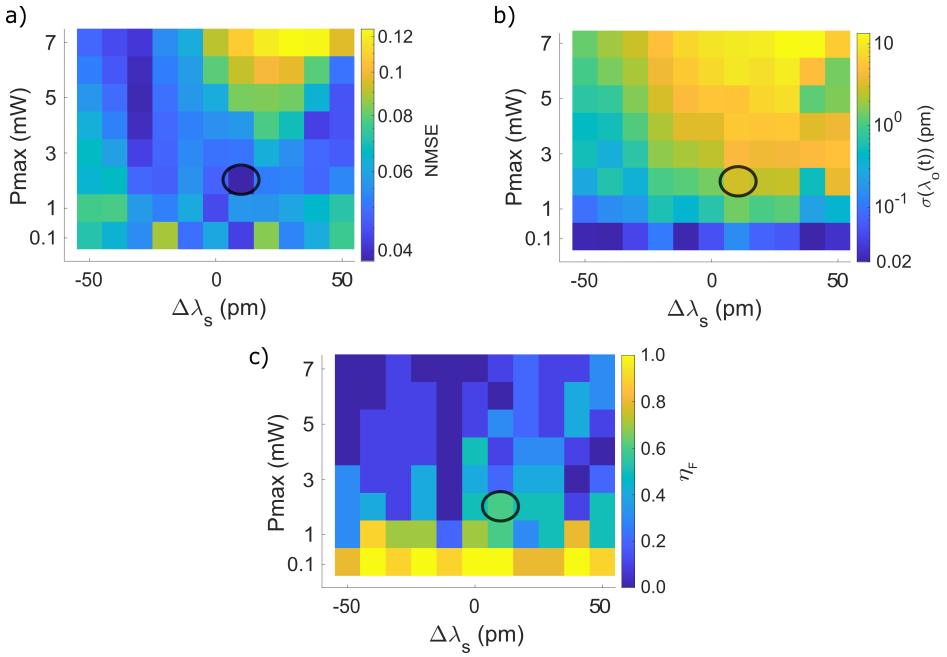

Figure 4.11: Performances relative to the Santa Fe benchmark task, by using a single microring with external feedback. (a) NMSE, (b) standard deviation of the resonance wavelength shift $\sigma(\lambda_0(t))$, and (c) strength of the feedback, versus the starting detuning $\Delta\lambda_s$ and the maximum incident power P_{max}. The quantities refer to the feedback configuration with the lower error achieved at each ($\Delta\lambda_s$, P_{max}). The black circle indicates the conditions for which the lowest NMSE is obtained.

maximized (Fig. 4.11(c), at $P_{max} = 0.1mW$). Thus, the recursivity of the feedback signal in the system is increased and, consequently, the detection nonlinearity acts on a larger number of feedback-delayed terms. This suggests that when the microring is forced in a linear regime, due to the limited input optical power, the system enhances the linear memory using higher feedback strengths, to effectively solve the task. The best performance for this processing scheme results in $NMSE = 0.042 \pm 0.008$. At higher maximum incident power $P_{max} > 0.1mW$, the microring nonlinearity becomes also accessible, but is not required. As illustrated in the previous task, the feedback phase $\Delta\phi_F$ can still be tuned to minimize the optical power inside the microring (destructive interference), and thus minimize its nonlinearity. Nevertheless, Fig. 4.11(b) shows that for these input optical powers, the

standard deviation of the resonance shift, $\sigma(\lambda_0)$, is higher than in the linear case (Fig. 4.11(b), at $P_{max} = 0.1mW$), so that the system is actually exploiting the microring nonlinearity. In particular, a region of lower NMSEs in Fig. 4.11(a) is related to intermediate values of $\sigma(\lambda_0)$. In parallel, the strength of the feedback is reduced with respect to the linear case (Fig. 4.11(c)). These results suggest that, under these conditions, the microring contributes to the overall nonlinearity of the system and improves the prediction performance, down to $NMSE = 0.038 \pm 0.008$, at $\Delta\lambda_s = 10$ pm, $P_{max} = 2$ mW, $\eta_F = 0.55$ and $\Delta\phi_F = 0$. Alternatively, the detection nonlinearity is also sufficient to solve the task, once enhanced by stronger external optical feedback.

Interestingly, a better performance, compared to the linear shift register, is obtained when disconnecting the feedback from the microring ($\eta_F = 0$). The results are shown in Fig. 4.12, as a function of $\Delta\lambda_s$ and P_{max}. Figure 4.12(a) indicates the NMSE achieved, 4.12(b) the nonlinear microring resonance shift, and 4.12(c) an estimate of the memory capacity (MC) computed on the SantaFe input series itself, as previously discussed in section 4.3. The lowest obtained error is $NMSE_{min} = 0.045 \pm 0.002$, and it is achieved for $P_{max} = 7$ mW and $\Delta\lambda_s = -40$ pm (Fig. 4.12(a), black circle). This NMSE value is only slightly higher than the one obtained with a microring with an external loop, indicating that the contribution of the extended external memory is not critical for this task. When the microring works in the linear regime (i.e. for low optical power $P_{max} = 0.1$ mW), the NMSE obtained is as high as 0.13 (Fig. 4.12(a)). In this case, the microring has only access to an inertia memory of one step, as explained in section 4.3.1, while the output undergoes a nonlinear transformation through the square-law photodetection. Consequently, one bit of memory and nonlinearity are sufficient to obtain a NMSE lower than $NMSE_{SR}$. By increasing P_{max} the nonlinear contribution of the microring is increased and the $NMSE$ is significantly reduced (Fig. 4.12(a)). This happens only for those $\Delta\lambda_s$ conditions that preserve the free carrier nonlinearity of the microring in a specific range of values (Fig. 4.12(b)). The $\sigma(\lambda_o)$ should not be very low - which implies an absence of nonlinear effects - but also not high enough to trigger self-pulsing dynamics, as explained in section 4.3.2. The configurations leading to lower NMSE values are also correlated to an increased memory, with respect to the one that the microring has in the linear regime ($Pmax = 0.1$ mW). This difference is indicated as ΔMC in Fig. 4.12(c). For conditions that lead to self-pulsation dynamics (e.g. for high P_{max}, detuning dependent), the capability of the system to retain the memory is lower, the nonlinearity of

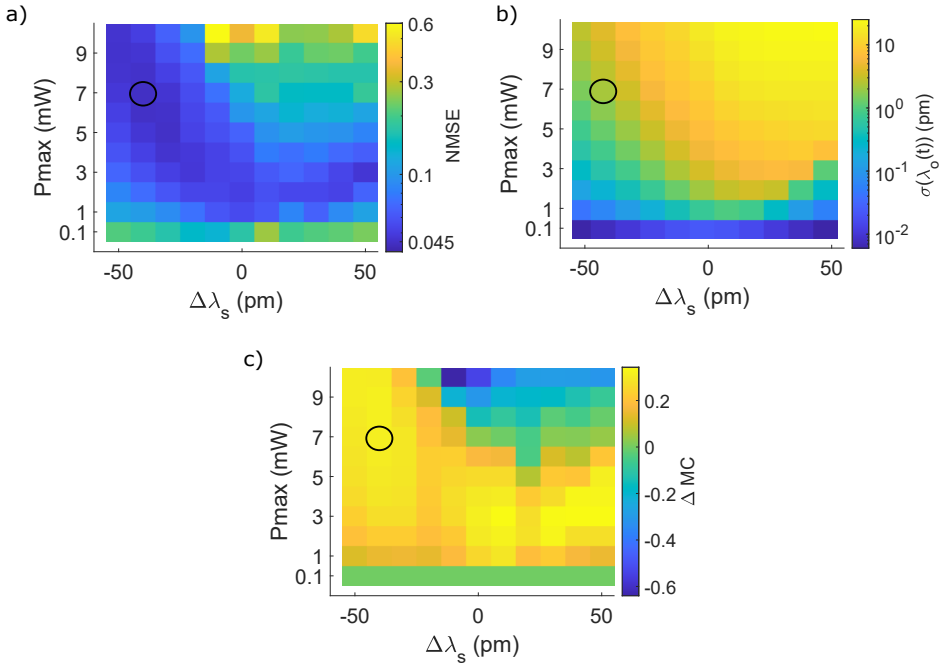

Figure 4.12: Performance of the Santa Fe benchmark task, by using a single microring without external feedback. (a) NMSE, (b) standard deviation of the resonance wavelength shift $\sigma(\lambda_0)$, and (c) change of MC (ΔMC), versus the starting wavelength shift $\Delta\lambda_s$ and the optical power P_{max}. The black circle indicates the conditions for which the lowest NMSE is obtained.

the system is very high, and the overall capability for computation becomes limited.

4.4 Noise effects on performance

Several noise sources can be present in the different layers of a computing system. Instruments at the input layer that encode the input information introduce noise, as well as those at the output layer that readout the node response. The reservoir itself can introduce noise, in particular when including active elements, such as optical amplifiers. The presence of noise can be beneficial or a deteriorating factor for the computational properties of a system. For instance, in [125] noise is intentionally added to the input series

(x_i). The network can in this way dynamically explore a larger region in the state space, which results, after the training phase, in more robust long-term predictions. In contrast, the presence of noise at the output layer may have negative effects on the performance, with an impact that also depends on the task: classification tasks are noise-robust tasks, while memory and prediction tasks, where the target is a precise analog value, may result more sensitive to noise [126].

In this section, the role that the microring nonlinearity has in preserving the task performance when noise is introduced at the output layer, is investigated [127]. To do so, the single microring resonator, in absence of feedback, is tested on the Santa Fe benchmark task. Here, Gaussian white noise is added to the drop response of the system while solving the task as described by the following equation:

$$|E_{drop}(t)|^2 = |E_{drop}(t)|^2 + 10^{(noise(dB)/10)} Cu(t), \tag{4.15}$$

where $C = 0.16$ mW is a reference value, corresponding to the standard deviation of the detected drop response while processing the task in one of the configurations more resilient against noise (black circle in Fig. 4.14). u(t) is a random scalar value drawn from the standard normal distribution with mean 0 and variance 1.

The performance on the Santa Fe task for different amounts of output noise is reported in Fig. 4.13(a), versus the P_{max} and $\Delta\lambda_s$ parameters. The average noise value is indicated (in dB) at the top of each panel. In absence of output noise, the results are those already reported in the previous section 4.3.3, with a region of optimal performances clearly present, which extends both at positive and negative $\Delta\lambda_s$. The best configuration is achieved at $P_{max} = 7$ mW and $\Delta\lambda_s = -40$ pm. As the average noise applied on the drop signal becomes stronger, from -35 dB to -12 dB with respect to the reference value C, the computed error increases. The region with the best performance slightly moves at positive starting detuning. This is better displayed in Fig. 4.13(b), where the NMSE against the noise is shown for two particular configurations: the best one in absence of output noise ($P_{max} = 7$ mW, $\Delta\lambda_s = -40$), and one with positive starting detuning which results more resilient against noise ($P_{max} = 3mW$, $\Delta\lambda_s = 40$ pm).

More insights regarding this diverse noise resistance can be grasped from Fig. 4.14. In Fig. 4.14(a), it is shown the standard deviation of the resonance shift while computing the task, $\sigma(\Delta\lambda_o(t))$, for each scanned ($P_{max}$,

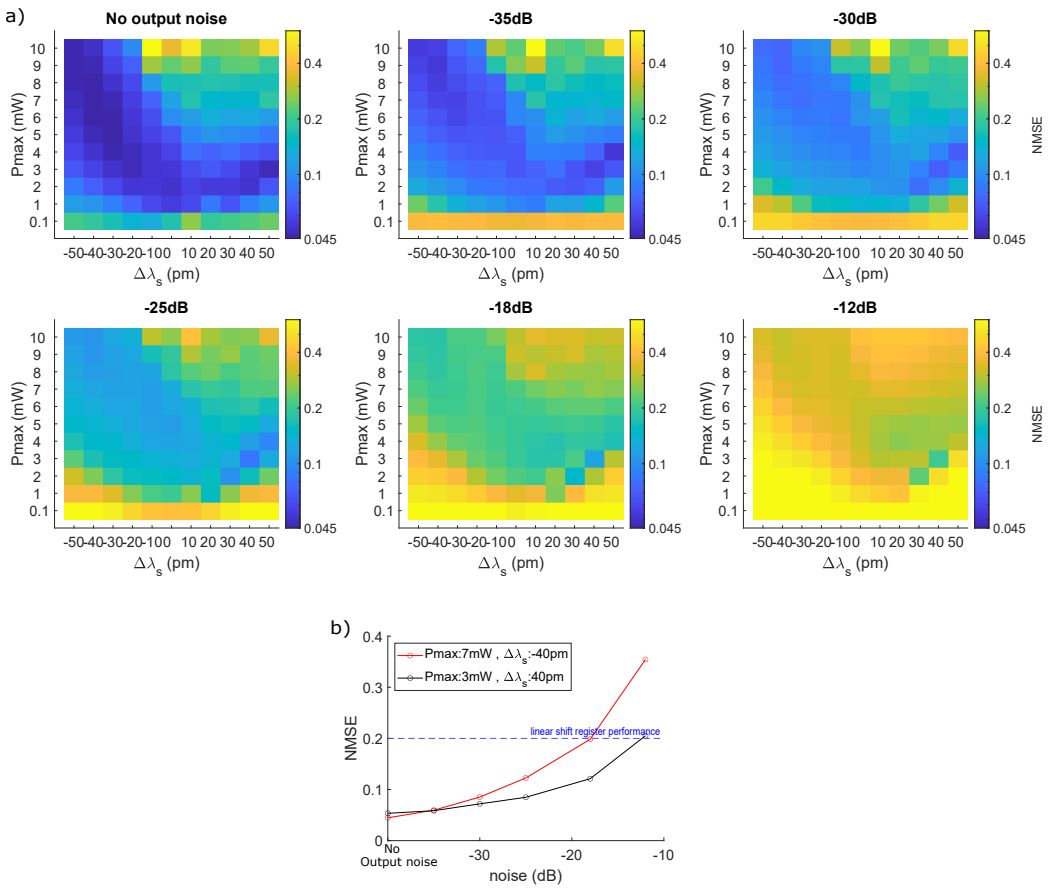

Figure 4.13: Performance on the Santa Fe benchmark task when using a single MRR in the add-drop filter configuration. a) NMSE vs the input optical power P_{max} and starting detuning $\Delta\lambda_s$, for different levels of noise, reported at the top of each panel. The noise is added directly at the drop signal. b) NMSE versus the amount of noise for 2 particular configurations: the configuration having lower NMSE in absence of output noise ($P_{max} = 7mW$, $\Delta\lambda_s = -40$ pm, red line), and a configuration which is more noise robust ($P_{max} = 3mW$, $\Delta\lambda_s = +40$ pm, black line). The blue dotted line indicates a reference value for NMSE, achieved by processing the input series x_i directly with a linear shift register, thus in absence of output noise.

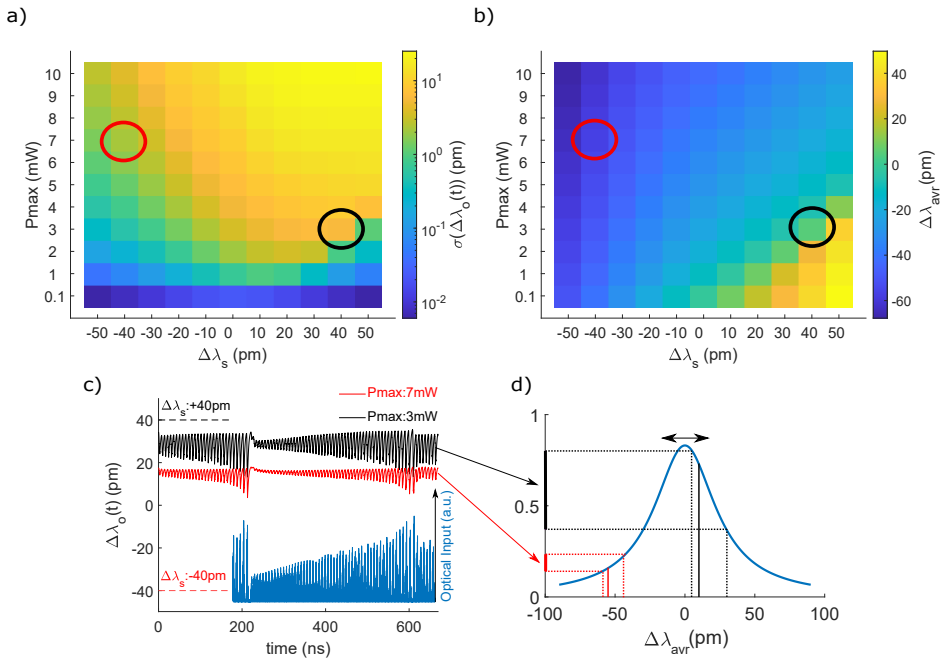

Figure 4.14: a) Standard deviation of the resonance wavelength shift $\sigma(\Delta\lambda_o(t))$ while computing the Santa Fe task, for each (P_{max}, $\Delta\lambda_s$) configuration, and b) effective average detuning during computation. c) Temporal evolution of the resonance wavelength shift to the Santa Fe input optical power (blue line) for two configurations: the one having lower NMSE in absence of noise ($P_{max} = 7mW$, $\Delta\lambda_s = -40$ pm, red circle in Fig. 4.14(a),(b)), and a configuration which is more noise robust ($P_{max} = 3mW$, $\Delta\lambda_s = +40$ pm, the black circle in Fig. 4.14(a),(b)). Dashed lines show the correspondent starting detuning. d) Portion of the resonance curve transmission effectively explored while computing the task for the previous two configurations. Solid lines indicate the effective average detuning and the dashed line indicates the correspondent maximum and minimum detuning while computing the task.

$\Delta\lambda_s$) configuration. The plots of Fig. 4.14(a) appear asymmetric with respect to $\Delta\lambda_s = 0$, because of thermal effects. Even though slow with respect to the input modulation, the temperature variations, due to the generated free carriers, contribute with a positive bias to the resonance shift $\Delta\lambda_o(t)$. As a result, if initially $\lambda_p > \lambda_0$, the detuning decreases because of the temperature bias shift. As a consequence, the same optical power can be loaded within the microring using lower P_{max} values, as compared to negative detuning configurations, and induce the same nonlinear resonance shift $\sigma(\Delta\lambda_o(t))$. This effective average detuning can be expressed as $\Delta\lambda_{avr} = \Delta\lambda_s - \overline{\Delta\lambda_o(t)}$, where $\overline{\Delta\lambda_o(t)}$ is the average resonance shift. $\Delta\lambda_{avr}$ is the quantity plotted in Fig. 4.14(b). Interestingly, the configurations at positive starting detuning in Fig. 4.13(a), which are more robust against noise, share in Fig. 4.14(b) the same amount of $\Delta\lambda_{avr} \approx 10$ pm.

In Fig. 4.14(c) the resonance shift dynamics ($\Delta\lambda_o(t)$) is reported, for the lowest NMSE configuration in absence of noise ($P_{max} = 7mW$, $\Delta\lambda_s = -40$ pm, red circle in Fig. 4.14(a),(b)) and for a more noise-robust configuration ($P_{max} = 3mW$, $\Delta\lambda_s = +40$ pm, black circle in Fig. 4.14(a),(b)). The blue line refers to the optical input to the microring in arbitrary units, while the dotted line refers to the starting detuning $\Delta\lambda_s$ for the two configurations. In both cases, the resonance oscillates in phase with the optical input. But in the more noise-robust configuration (black line), the effective average detuning is lower, as the resonance positions oscillate closer to the pump wavelength ($\Delta\lambda_s = 40$ pm). These facts are mapped on the resonance transmission shape in Fig. 4.14(d) for clarity. The noise-robust configuration not only shows a larger resonance oscillation range, but it also corresponds to a steeper region of the resonance curve. As a result, the same resonance shift imprints more nonlinear effects on the drop signal, than the same resonance shift interesting a smoother region of the resonance curve, i.e. that for the red configuration. Thus, the configuration which is more noise-robust is actually the one that exploits more the microring nonlinear transformation.
Similar conclusions are also found in [128], where an electronic implementation of time-delay RC was studied. An electronic real node applies a Mackey-Glass nonlinear transformation to the incoming signal, and the memory of the system only relies on a feedback signal. With that system, noise-robust system configurations are associated to the nonlinear region of the Mackey-Glass function that is explored during the Santa Fe task. Nevertheless, exploring the highest nonlinearity region of the function is not helpful for computation: the delay-based network has in this case a limited linear memory capacity

and solves the task with worse performance. On the contrary, the microring studied here does not have a feedback signal and the only source of memory comes from the microring nonlinearity inertia. Thus, configurations with higher nonlinearities improve the nonlinear memory, and eventually the performance, as long as thermal effects are not so important (for large P_{max}).

Chapter 5

Fiber-based experimental implementation

In this section an experimental realization of a time-delay RC, based on the dynamics of a silicon microring resonator coupled to an external optical feedback, is presented. The design of the experiment relies on the numerical results obtained in chapter 4, with particular attention to replicating the system parameters that emerged critical for computation. These are the average input power, the starting detuning, the strength, and the phase of the feedback signal. This physical implementation involves a microring resonator designed in an add-drop filter configuration on a SOI chip, with a grating coupler at each of its ports for external accessing the structure. The feedback is realized by connecting the through and add ports, like in the numerical study, via an external single-mode optical fiber.

The chapter is organized as follows. First, the interferometric problems that an optical fiber-based feedback introduces in the system are discussed, together with a proposed solution. Then, the experimental setup for testing the microring-based RC computational capabilities is described, together with an experiment aiming to define the fading memory of the structure.

5.1 Hybrid approach

An integrated microring coupled to an external optical fiber may refer to a hybrid approach of time-delay RC, where the system is only partially integrated. A fully integrated version of the feedback has been also designed,

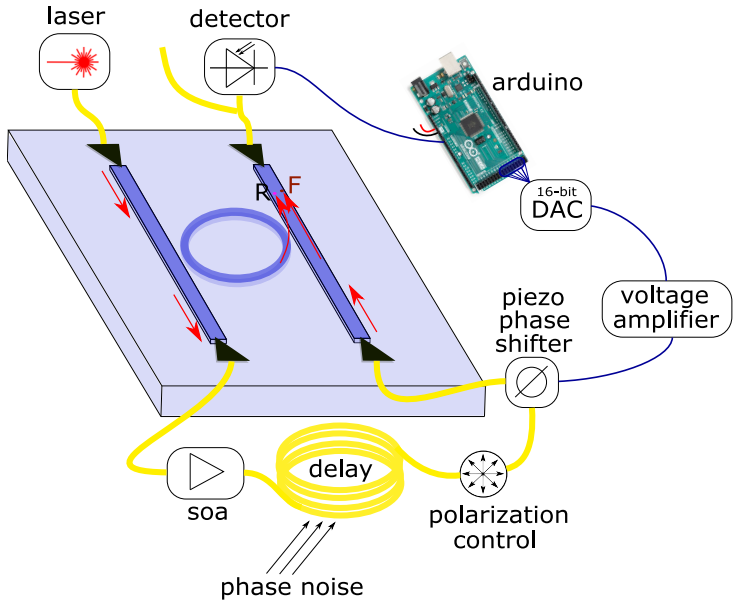

Figure 5.1: Time delay RC based on a microring and an optical fiber feedback loop. The optical fiber connects the through and add ports. An Arduino-based controller stabilizes the system against environmental noise. 'R' and 'F' refer to the optical contributions to the drop signal, from the ring and the feedback respectively, respectively. Along the feedback loop there are a semiconductor optical amplifier SOA, a polarization controller, and a piezo-phase shifter.

as will be discussed in section 7.1. When the feedback is realized with an optical fiber, the optical signal propagating through the fiber is affected by environmental thermal, phonic, and vibrational noise. All these noise sources change the refractive index of the fiber and thus the phase of the delayed signal. These changes can interest all the fiber or only a portion of it, depending on where the perturbation occurs. An increase in the room temperature or people speaking loudly close to the fiber, for example, perturb its length. On the other hand, touching the fiber in a particular region only affects locally the refractive index. In both cases, the phase of the feedback signal is varied over time, with a speed that depends both on the noise source and on the interested fiber length. As a consequence, when the feedback signal is coupled back to the microring, noisy interference conditions take place.
The drop signal is the result of the interference between the optical field from

the microring (R) and the field from the feedback loop (F), as represented in Fig. 5.1. These two fields can be expressed in a simple way that highlights their relative phase as $R = A$ and $F = Be^{i\Delta\phi_F}$. As a result, the optical power at the drop port is:

$$|R + F|^2 = A^2 + B^2 + 2ABcos(\Delta\phi_F), \tag{5.1}$$

and ranges between maximum and minimum identified by $A^2 + B^2 \pm 2AB$. The amplitude of the two field components defines the interference range, while the relative phase states the interference condition: constructive ($\Delta\phi_F = 0$), destructive ($\Delta\phi_F = \pi$), or intermediate, with π-periodicity. Environmental noise randomly changes the phase of the delayed signal and is thus computed with Eq. 5.1, as fluctuations in the $\Delta\phi_F$ values over time. As a result, the drop signal, which is driven by this noise, randomly varies within the interference intensity range, even when a CW input signal is injected. This problem was not taken into account during the modeling work, where the feedback phase was not affected by any perturbation, and was a real challenge to face for the practical realization of the time-delay RC.

5.2 Phase stabilization

This section describes the system that has been developed to compensate for the noisy phase fluctuations of the feedback signal in the microring system with external fiber loop. The system is inspired by the working principle of a Proportional Integrative Derivative (PID) controller, a feedback control mechanism widely used in a variety of applications requiring continuously modulated control over a physical quantity. Here, after fixing a desired target value for this quantity, the PID controller continuously calculates an error value $err(t)$ based on the difference between the target and its current value. The controller applies then a correction based on proportional, integrative, and derivative terms, with the last two terms accounting for the response time of the system and the rate at which the physical quantity of interest is changing.

5.2.1 The phase controller

To compensate for environmental noise, a specific controller has been developed. Similarly to a PID controller, this system needs to detect the effects

of the noisy fluctuations affecting the microring-based system and continuously compensate them according to a target value. The phase controller developed is composed of the following instruments:

- Detector: placed at the drop port, this detector measures the power fluctuations induced by phase noise along the feedback loop.
- Arduino Mega: computing unit, which continuously estimates a correction signal aiming to keep the phase of the feedback loop at a desired target value.
- Digital Analog Converter (DAC): it converts the digital correction signal from Arduino into an analog electronic signal.
- Voltage amplifier: it amplifies the electronic correction signal.
- piezo-driven phase shifter (LUNA FPS-001): this instrument is only a few centimeters long, and has an optical fiber connected to its ends. The driving voltage allows to change in the length of the piezo material, hence the length of the internal optical fiber and, as a result, also the optical path and the phase of the propagating light after propagating through the phase shifter.

A dedicated detector is placed at the drop port of the microring to measure the time varying drop signal due to the phase noise. This signal is then sent to an Arduino Mega, where it is stored as v_{in}. Arduino continuously elaborates a correction signal v_{out} based on the desired target value v_{tar}. The target value corresponds to a power value within the interference range and is set via PC using a serial connection. The correction signal v_{out} is encoded as a 16-bit integer and then transmitted to a DAC using 16 output logical channels available in the Arduino board. The resulting analog signal is amplified and then fed to a piezo-driven phase shifter (LUNA FPS-001, 0-20KHz bandwidth) placed along the feedback fiber line, which finally applies the phase correction. When the piezo-phase shifter is properly driven, phase noise fluctuations are compensated. The process has to run fast enough to correct the noisy fluctuations. Arduino Mega benefits from a 16 MHz clock speed, that is used to run the list of instructions needed for each correction action. This list is reported in Code 1 in appendix C. The program updates the output every $55\mu s$, which accounts for the external digital to analog conversion, as shown in Fig. 5.2. The voltage amplifier, with only 200 Hz bandwidth, limits

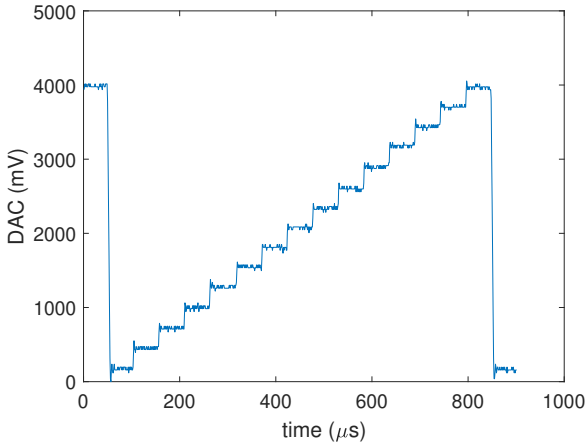

Figure 5.2: DAC output voltage, when driven by Arduino running the phase correction code 1 in the absence of any input. New outputs are provided every $55\mu s$ approximately, thus at $18kbps$, and the maximum output voltage at the DAC is $\sim 4V$. The value of the voltage steps shown is a consequence of the proportional constant (kp) set.

the bandwidth of the entire phase controller system. Still, the system is able to correct the largest contribution of the environmental phase noise, which is present at lower frequencies, as will be discussed in section 5.2.2.

Note that only the proportional component k_p is used in Code 1 when updating v_{out}. Indeed, $v_{out} = v_{out} + k_p(v_{tar} - v_{in})$. The phase of the feedback optical signal reacts so quickly with respect to the external perturbations, that is actually independent of past phase values. Thus, there is no need to use integrative and derivative terms when estimating the correction signal v_{out}. Note also that if this was not the case, and past states needed to be considered, a dedicated vector would have been needed to be stored on Arduino, with a size proportional to the inertia of the system, and updated at each void loop repetition, strongly limiting the correction rate.

5.2.2 Results

The effects of the noise at the microring drop port and its compensation using the phase controller system can be quantified by introducing into the system a CW optical signal and performing a Fourier transform of the drop

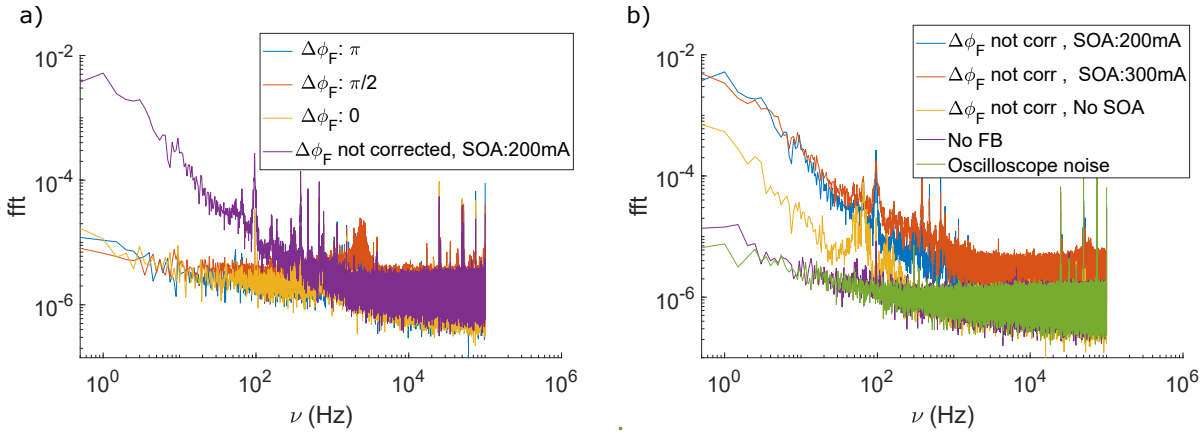

Figure 5.3: Fourier transform of the temporal traces acquired at -4GHz detuning from the microring resonance. a) The SOA is switched ON, while the stabilization controller is both OFF (purple line) or active at different phase targets. b) The stabilization controller is switched OFF, and the effects of the presence of the feedback are reported for various SOA configurations, including when the SOA is removed from the feedback loop (No SOA). As a reference, the noisy trace recorded by a detector with no input signal is also reported (detector and oscilloscope noise).

signal. The input optical signal needs to be partially coupled to the microring resonator in order to allow and probe variations in the interference taking place at the drop port.

The drop noise spectra are reported in Fig. 5.3 when the phase-controller is on/off, respectively. Different operational conditions are considered. When the phase controller is not active, and hence $\Delta\phi_F$ is subject to phase noise, the frequency contributions to the optical signal are larger at lower frequencies and decrease at higher frequencies. Fig. 5.3(a) shows that by properly driving the piezo phase shifter along the feedback loop with the phase controller, it is possible to substantially reduce the noise fluctuations at lower frequencies, up to approximately $1kHz$. Furthermore, Fig. 5.3(b) shows that many peaks at frequencies $\nu > 10^4$ Hz refer to artifacts due to noise in the oscilloscope and in the detector, rather than optical variations of the drop signal. Differently, the peak at 100 Hz is due to the power supply of some instruments, such as the biasing controllers of the laser and SOA, which is reflected in power variations at the drop port and amplified by the presence

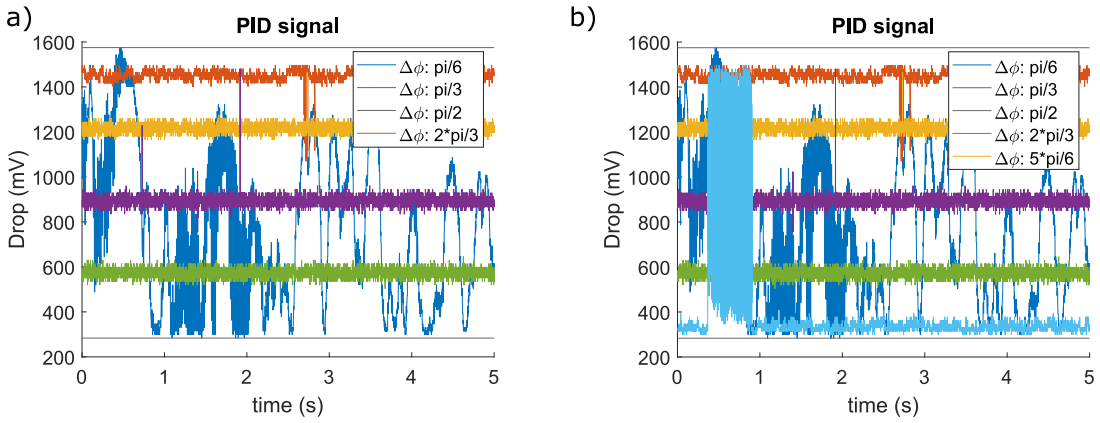

Figure 5.4: Microring drop response to a CW input. Data are taken by a 3 kHz bandwidth photodetector and an oscilloscope with a 1 ms sampling time. a) The blue curve in the background represents the noisy signal induced by environmental noise, while the colored lines represent the stabilized signal via the Arduino-based controller, at interference conditions specified by the user. b) Sometimes the correction fails, in particular when the controller target is close to the interference limits, or a strong perturbation occurs.

of the feedback. This peak highlights a potential source of error suffered by the phase controller: in the frequency range where the controller works ($< 1kHz$), any optical power variation at the drop port is compensated by varying the phase of the optical signal along the feedback loop, even though they may be induced by other factors, as for example by power supply fluctuations of some instruments. The phase controller attempts to compensate them, thus leading to wrong phase corrections. Luckily, the contribution at 100 Hz is a hundred to thousand times smaller than those induced by phase noise at lower frequencies and can be neglected.

Fig. 5.4 shows the temporal variation of the output signal over a five seconds time interval. It can be observed that the controller compensates the phase fluctuations over time. The correction may also fail, as shown in Fig. 5.4(b), in particular when the controller target is close to the interference limits, or a strong perturbation occurs. I anticipate here that the benchmark tasks that will be described in chapter 6.2, require about $300\mu s$ to be processed optically. This time is relatively fast compared to the phase noise fluctuations. Still, the phase controller is fundamental because allows setting the desired feedback phase value, maintaining it among several repetitions

of the task measures. Notably, from Fig. 5.4, we deduce a $\sigma_{\pi/6} = 17.5mV$, $\sigma_{\pi/3} = 20.4mV$, $\sigma_{\pi/2} = 20.1mV$, $\sigma_{2\pi/3} = 20.5mV$, and $\sigma_{5\pi/6} = 16.7mV$ (without considering the temporal interval where the controller fails), for the normalized output signal.

5.3 Experimental setup

The setup composed to investigate the time-delay RC is reported in Fig. 5.5. It can be divided into four parts: the input layer, the Device Under Test (DUT), the output layer, and the controller developed for phase stabilization.

The input layer is dedicated to the encoding of data in the optical domain and their routing in the DUT. First, light is generated by a Tunable Laser Source (TLS) operating in CW within the C-band. The optical signal is then amplitude modulated by using a mach zehnder interferometer (MZI) electro-optic IQ modulator (MXAN-LN-10, 10 GHz), driven by a 500MHz Teledyne (T3AFG500) arbitrary waveform generator (AWG), and stabilized in quadrature by a bias controller (MBC-AN-LAB). Then, an amplification step is performed by an erbium-doped fiber amplifier (EDFA), followed by a variable optical attenuator (VOA) for tuning the input power. A polarizer allows to proper polarize the input optical signal to match the TE polarization required by the DUT. Before entering the DUT, 1% of the input optical power is extracted by a directional coupler and routed towards a photodetector (PM_i), for monitoring the average input optical power. The modulated optical signal is input in the DUT, whose detailed description is given in section 5.3.1. At the output of the DUT, an optical filter suppresses amplified spontaneous emissions (ASE) noise generated by the SOA present in the setup. Then, after the optical filter, the signal is separated along two channels. One is labeled as the output layer of the system, and includes those instruments that allow acquiring the optical response of the microring to a certain task input series. This channel receives 90% of the DUT optical response. The optical signal is firstly photodetected (PM_t, Menlosystem, FPD610-FC-NIR, 600 MHz), then read out by an oscilloscope (Picoscope 6000 series, full bandwidth of 500 MHz), and finally communicated to a computer where it is stored for an offline evaluation. Note that the bandwidths of both the detector PM_t (600 MHz) and the picoscope (500 MHz), are chosen such to make the output layer able to detect the DUT response even for the largest AWG modulation frequency (500 MHz). The second channel aims to

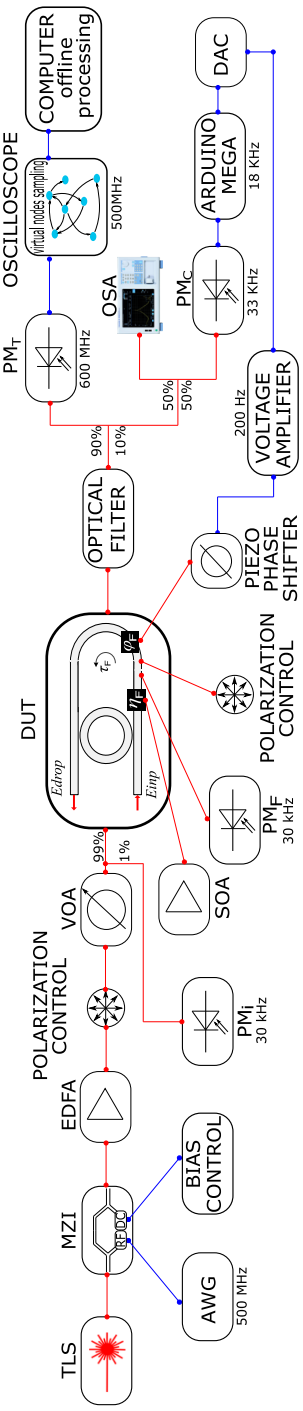

Figure 5.5: Scheme of the experimental setup.

compensate for environmental phase variations and involves the controller described in Section 5.2. Thus, 1% of the DUT response is directed to this channel using a directional coupler. Here, the DUT response is measured by means of a photodetector (PM_c) that is slower than the one employed in the output task layer. Its electronic signal is then input to the Arduino-based controller for phase stabilization. It is worth noting the importance of a slower detector in this channel. The detector PM_c needs to capture only the DUT optical response variations induced by phase environmental noise, and not the optical modulation that encodes the information to be processed. Indeed, the phase controller interprets any optical variation at the drop port of the microring as induced by environmental phase noise. In this way, any optical variation due to the input task modulation, if input to the phase controller, would lead to wrong phase corrections. Thus, a proper detector bandwidth is needed, possibly lower than the task modulation speed but higher than the main frequency contributions to environmental phase noise ($< kHz$). To this purpose, an InGaAs switchable gain detector (PDA10CS2, Thorlabs) is operated, whose bandwidth depends on the electronic gain set. Practically, the detector was set at a 70 dB gain, with a consequent 3 kHz bandwidth (at 1550nm), sufficient for detecting most of the phase noise, and too slow for the task modulation rate, that, as will be discussed later on, is $\sim 20MHz$.

5.3.1 DUT: microring resonator with feedback loop

The device under test (DUT) corresponds to the microring resonator and the feedback loop, as sketched in Fig. 5.6. The microring is fabricated on a SOI chip, in an add-drop filter configuration. All four ports of the microring are accessible by means of grating couplers (GCs), which are aligned and equally spaced by $127\mu m$, for coupling an 8-channel SM fiber array, with 6.9° polishing angle. In this way, all the ports are accessible at the same time, in agreement with the numerical model studied in section 4.2.3. The fiber array is fixed on a fiber holder, which in turn is moved with sub-micrometric precision along the x, y, and z directions, thanks to a Thorlabs 3-axis NanoMax 300 stage. The alignment between the fiber array channels and the GCs is optimized, with each GC having $5dB$ coupling losses. The silicon-on-insulator chip is fixed on top of an aluminum block, which is positioned on a Peltier cell. The aluminum block, however, was not equipped with a hole for a temperature sensor, so the chip was not thermalized. Thus, to compensate for

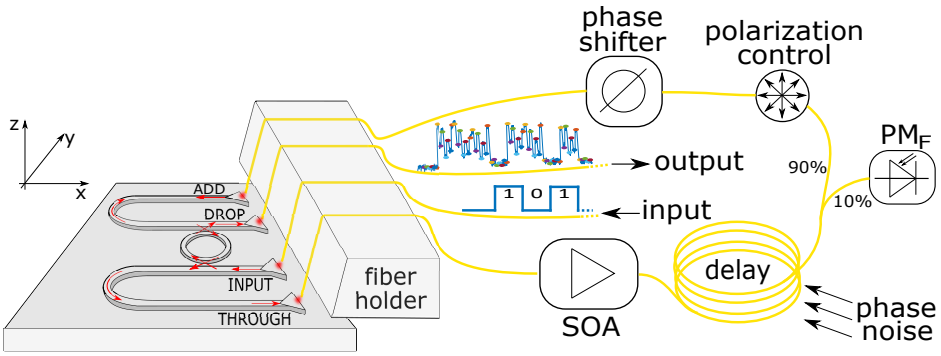

Figure 5.6: Experimental device under test: microring resonator with external optical feedback. The microring is integrated in a silicon-on-insulator chip and accessed by a fiber array. The feedback delay line is realized externally, via an SM optical fiber. It includes a series of instruments that allow to control and monitor the phase, the strength, and the polarization of the feedback signal. The input signal comes from the input layer, while the output signal is sent to the output optical fiber.

environmental thermal fluctuations of the microring resonance position, this last is tracked using automatized routines during a measure. The measures are moreover automatized and run overnight when the laboratory temperature is more stable, achieving resonance deviations, with respect to the starting value, lower than 1 GHz.

An external single-mode optical fiber is coupled to the through and add ports of the microring, providing the feedback loop for the application. Along the feedback loop, many instruments are used to manipulate and monitor the optical feedback signal. An SOA is used to vary the strength of the feedback signal. Its driving electric current ranges between $[0 - 500]mA$, and determines whether the SOA works as a gain medium, increasing the strength of the feedback signal, or as an absorber medium, decreasing the strength of the feedback signal. 10% of the feedback signal is extracted and routed toward a monitor detector. The optical signal is polarized to match the TE polarization, and then phase modulated by a piezo-phase shifter, to set the desired phase value while compensating for the environmental phase noise. Finally, the feedback signal is coupled back to the chip.

Note that the delay provided by the feedback loop is mostly induced by the propagation in the SM fiber. However, also the other components in the

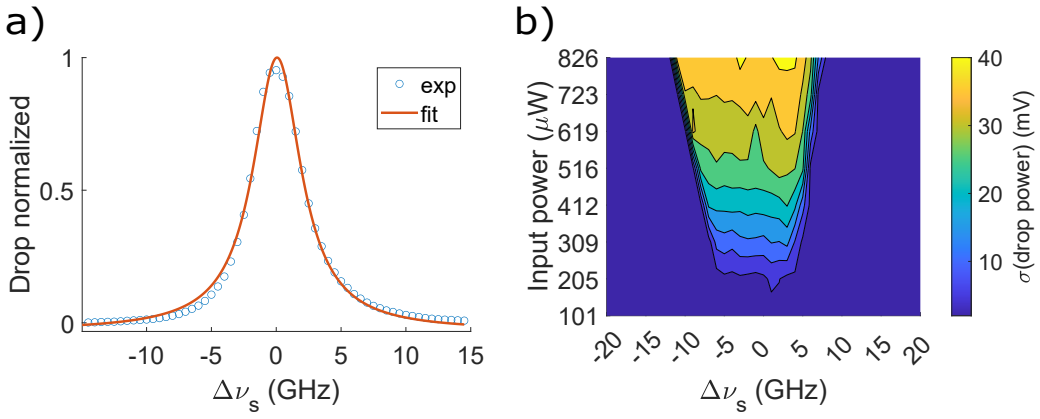

Figure 5.7: Microring resonator in absence of feedback: a) Drop transmission and b) self pulsing map, evaluated as the standard deviation of temporal trace, for different configuration of detuning ($\Delta\nu_s$) and optical input power (P_{in}).

delay line contribute to the total feedback delay time. The SOA, internally, is coupled to an optical fiber of approximately $2m$ long; the fiber array also introduces delays due to the length of each optical fiber channel. The total feedback delay time can be estimated without prior knowledge of the delays introduced by the instruments in the delay line. In section 5.5 a simple experiment to measure the total delay is presented.

5.3.2 Microring resonator characteristics

The integrated microring used in the experiment has a waveguide cross-section of $220 \times 500nm^2$, a radius $r = 6.75\mu m$, and a symmetric gap with the two bus waveguides of $238nm$. Its drop port transmission is reported in Fig. 5.7(a). From this spectrum we deduce a $FWHM = 4.52GHz$ ($FWHM = 36.4pm$), and a quality factor $Q = 4.2 \times 10^4$, at a resonance cold position $\nu_0 = 193.16THz$ ($\lambda_0 = 1553.11nm$).
At high input optical power, the microring enters in self-pulsation dynamics. This can be observed by plotting the standard deviation of the drop temporal traces to CW input optical signals at different configurations of detuning ($\Delta\nu_s$) and input power (P_{in}). As can be observed from Fig. 5.7(b), for $P_{in} > 200\mu W$, the ring starts to exhibit oscillations that become more important as the power is increased. Note also that larger $|\Delta\nu_s|$ values re-

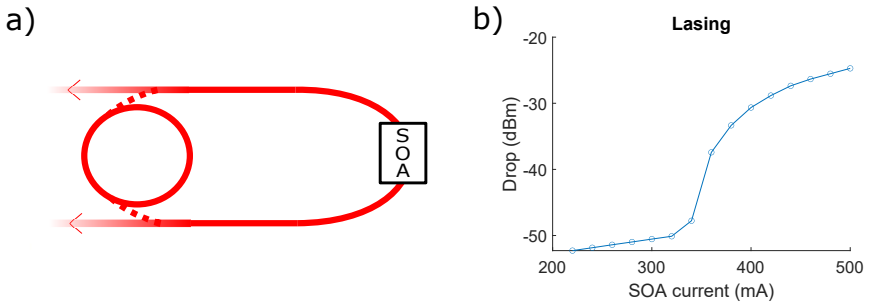

Figure 5.8: a) Lasing circuit and b) Drop power at the microring resonance wavelength $\lambda_0 = 1553.11nm$ ($\nu_0 = 193.16THz$).

quire larger input power to initiate self-pulsing. The latter are finally not triggered when the detuning is completely out of resonance.

5.4 Lasing

An important aspect of the setup is the possibility of lasing in the feedback loop, due to the presence of the optical amplification section. The SOA emits amplified spontaneous emission noise over a wide range of wavelengths. Those wavelengths that are resonant with the microring are coupled back to the feedback SOA itself, hence producing a positive feedback loop (Fig. 5.8(a)).
Lasing requires the gain to be larger than the losses along the feedback loop. The SOA wavelengths are affected by different losses, for example because of the grating couplers frequency response. In the experiments that follow, lasing occurs only at the working cold resonance wavelength $\lambda_0 = 1553.11nm$, and can be observed just increasing the current driving the SOA while keeping the input laser off, as shown in Fig. 5.8(b).

Despite lasing was not considered during the numerical work, it represented another experimental challenge. Signatures of lasing emerged during the calibration of the phase controller, which was not able to stabilize the phase of the optical feedback signal.

5.5 Echo memory experiment

The fading memory is an important property in time-delay RC, representing the fact that the information stored in the system needs to fade over time. In this way, the dynamical response of the reservoir depends only on recently injected input.

In our system, the memory of previously injected bits can be encoded in an optical signal that iterates several times the feedback loop. The experiment that is here presented aims at measuring these multiple iterations. The experiment consists in measuring the drop response to a series of input optical pulses injected to the microring system. The pulses have a duration $bw = 10ns$ and are spaced by a time much larger than $\tau_F = 88ns$. In this way, for every input pulse, it is possible to observe if there are optical delayed pulses at the drop port.

The average input optical power to the microring is $P_{in} = 78\mu W$, which, according to Fig. 5.7, is a value well below the appearance of self-pulsing phenomena. The experiment is carried out for different configurations of detuning $\Delta\nu_s$, feedback strength η_F (SOA current), and feedback phase $\Delta\phi_F$. $\Delta\phi_F$ is set in the following way: for each combination of $\Delta\nu_s$ and η_F, the interference range at the drop port is firstly estimated by measuring the drop response fluctuations under environmental noise, using the detector PM_C for five seconds. Based on the maximum (constructive interference, $\Delta\phi_F = 0$) and minimum (destructive interference, $\Delta\phi_F = \pi$) values of this interference range, the target optical power related to the selected $\Delta\phi_F$ is estimated according to Eq. 5.1 and set on Arduino during the experiment. In particular, $\Delta\phi_F \in [\pi/6, \pi/3, \pi/2, 2\pi/3, 5\pi/6]$. The feedback strength is spanned, in terms of SOA current, in all the supported range $[0, 500]mA$, with steps of $50mA$. The detuning $\Delta\nu_s \in [-30, 30]GHz$, with steps of $6GHz$. Experimental traces acquired at the drop port are reported as an example in Fig. 5.9, for different combinations of detuning and SOA current, for a specific $\Delta\phi_F = 2\pi/3$. In this figure, every row is related to a detuning value, while every column is related to an SOA current. The first column shows the drop response in absence of any feedback (zero SOA current). In this condition, at the drop port, only the portion of the input pulse signal that is directly dropped can be observed, while the portion that, via the through port, is coupled to the feedback loop is fully absorbed by the SOA.

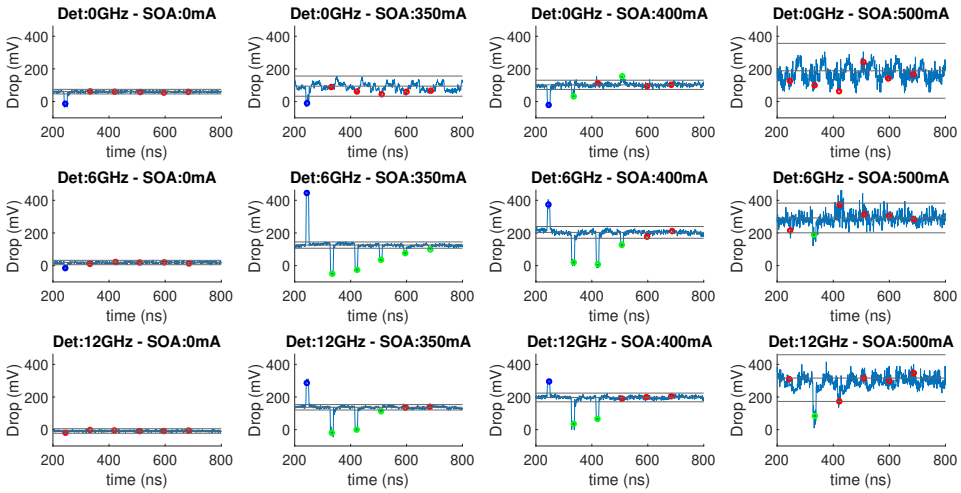

Figure 5.9: Drop response to a $10ns$ input pulse, acquired at $\Delta\phi_F = \pi/3$ and for various detuning (rows) and feedback strength η_F (SOA current) (columns). The first column represents the system response in absence of a feedback signal (zero SOA current). An increase in the SOA current allows a larger number of delayed copies (echoes) in the microring with external feedback. The blue dots indicate the drop copy that follows the input one, while the green dots refer to the copies via the feedback loop. The red dots indicate the delayed copies that are not counted by the algorithm.

Thus, only one optical pulse is observed at the drop port. Note that the height of dropped pulse decreases, as the detuning increases, since the input wavelength is moving away from the resonance condition. When η_F (SOA current) is increased, other pulses start to appear at the drop port. The latter represent the delayed copies of the input pulse that, thanks to the amplification step along the feedback loop, are able to iterate multiple times along the feedback circuit. In these cases, the first pulse observed at the drop port is provided by the portion of the input signal that is directly dropped, while the others are the delayed copies, all equally spaced by the delay time, which in this way can be estimated in $\tau_F = 88ns$. Note that the first dropped pulse, is practically simultaneous to the input pulse, if compared to τ_F. This contribution, during a task, will allow having the information of the current bit in the analysis of the virtual nodes. Information regarding the past processed bit will also contribute to the state of the virtual nodes,

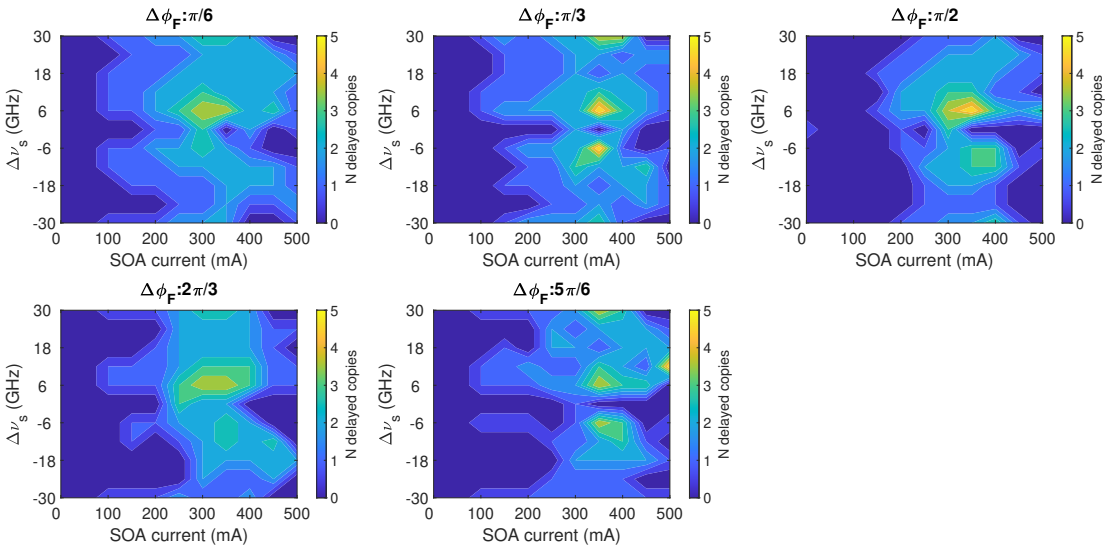

Figure 5.10: Number of delayed copies (green dots in Fig. 5.9) as a function of detuning and feedback SOA current, for five different values of $\Delta\phi_F$, ranging from almost constructive ($\Delta\phi_F = \pi/6$) to almost destructive interference ($\Delta\phi_F = 5\pi/6$. The number of delayed copies increases with η_F (SOA current), reaching a maximum of 5 delayed copies for SOA current toward $350mA$.

thanks to the delayed response of previous bits. Note, moreover, that starting from a feedback SOA current of $400mA$, the temporal traces become noisier, probably because of lasing dynamics (refer to Fig. 5.8).

The number of delayed copies related to each system configuration is reported in Fig. 5.10, for the 5 values of $\Delta\phi_F$ spanned. The number of delayed copies is estimated in the following way. Every temporal trace is acquired for $\approx 1300ns$, synchronously to a trigger signal, and averaged over three repetitions. The last $600ns$, characterized by an absence of delay copies, are exploited to estimate the mean value (P_{mean}) and the standard deviation ($\sigma(P)$) of the temporal trace in absence of pulses. A delayed copy at the drop port is effectively counted when $|P_{copy} - P_{mean}| > c * \sigma(P)$, being P_{copy} the optical power at the drop port, correspondent to the peak of the delayed copy, and c an adjustable parameter. As one can observe, the number of delayed copies increases when larger feedback SOA currents are used, reaching maximum values toward an SOA current of $350mA$ and detuning values of

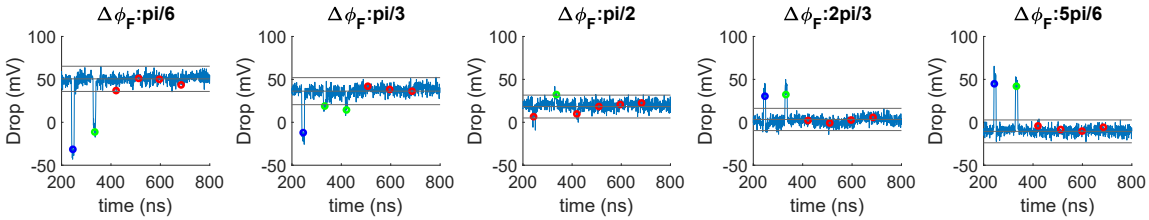

Figure 5.11: $\Delta\phi_F$ effects on the temporal traces while operating at $6GHz$ detuning and $150mA$ SOA current. The delayed copies tend to vanish while approaching $\Delta\phi_F = \pi/2$.

$\pm 6GHz$, with 5 delayed copies. Note however that this number represents an underestimation. A finer tuning of the SOA current may allow the observation of a higher number of delayed copies, before entering the lasing behavior.

Fig. 5.10 also shows that, for small feedback SOA currents, no delayed copies are measured. This region extends to higher η_F (SOA current), when $\Delta\phi_F \to \pi/2$. The absence of delayed copies at lower SOA currents is caused by the high optical losses in the feedback loop, mainly due to the two grating couplers, which are not sufficiently compensated by the SOA amplification. To understand why the region without delayed memory extends to larger SOA current when $\Delta\phi_F \to \pi/2$, the reader can refer to Fig. 5.11. As can be observed, the different $\Delta\phi_F$ values reflect different average optical power at the drop port. In the particular case $\Delta\phi_F = \pi/2$, the delayed copies tend to vanish.

Chapter 6

Experimental results

In this section I present the results that the microring with an external delay loop system has obtained in benchmark tasks. The tested tasks are in particular delayed boolean operations that require both linear and nonlinear memory to be solved. In this way, the computational properties of the system such as its nonlinearity, memory, and consistency over time are tested for time delay RC. The experiments are performed with an input optical power to the microring resonator well below self-pulsation, where the microring regime is linear. This maximizes the linear memory, while using the photodetection as the main nonlinearity source in the system. The same system configuration was exploited in the numerical work in section 4.3.1, to solve the Narma 10 task. The microring is coupled to an optical feedback loop where we allocated a maximum of 8 virtual nodes. The positive results obtained in the tested delayed-boolean tasks prove the consistency over time of the system, providing in particular indication regarding the successful implementation of the phase controller in the system. Nevertheless, the low number of virtual nodes and input optical power, limited the performances in time series prediction tasks, such as the Santa Fe and Mackey Glass tasks reported in the numerical work. These are thus omitted from the discussion and represent future benchmark tasks on which testing the system when the number of virtual nodes will be improved and a wider range of input optical power will be investigated.

The chapter is organised in this way. First, the chosen boolean tasks are described in section 6.1. Then, the results obtained on these tasks both including and excluding the feedback loop are discussed in section 6.2 and section 6.3, respectively. A final discussion where these two approaches are

compared concludes the chapter.

6.1 Tasks description

Time delay RC uses time multiplexing to serialize the information to process. Delayed boolean tasks refer to standard logical operations between two input bits, which are delayed in time because of time-multiplexing at the input layer of the system. The different boolean tasks investigated in this work are reported in table 6.1. They consider considered common logical tasks such as the AND, OR, NAND, and XOR. Also a linear memory task is introduced and labeled by MC, having the target always equal to the delayed bit value. Note that the XOR task is considered as the most difficult two-bit binary delayed task since it is not linearly separable in machine learning terms. The XOR task represents for this reason a nonlinear memory task. Each

x_i	x_{i-d}	MC	AND	NAND	OR	XOR
0	0	0	0	1	0	0
0	1	1	0	1	1	1
1	0	0	0	1	1	1
1	1	1	1	0	1	0

Table 6.1: Delayed-boolean tasks operated between the actual bit x_i and the previous d intervals in the past bit x_{i-d}.

boolean operation applies between the current bit x_i and a delayed bit x_{i-d}, with $d = 0, 1, 2, \ldots$, and has target y_i dictated by the logical operation considered.

The input series x_i is thus a pseudo-random binary sequence (PRBS) of bits, which is then time-multiplexed at the input layer of the network, in the optical domain. $x_i = 0$ does not necessarily correspond to zero optical power, as it depends on how the Mach-Zehnder modulator is operated. From the microring response, the output o_i is computed. Only the dynamical response to the optical input x_i is used for sampling the virtual nodes, as sketched in Fig. 6.1. In this way, the memory properties of the microring-feedback system are tested. In an alternative approach [106], virtual nodes are sampled from multiple b_w time intervals at the output response, which extend in the past direction. In this case, artificial memory is introduced at

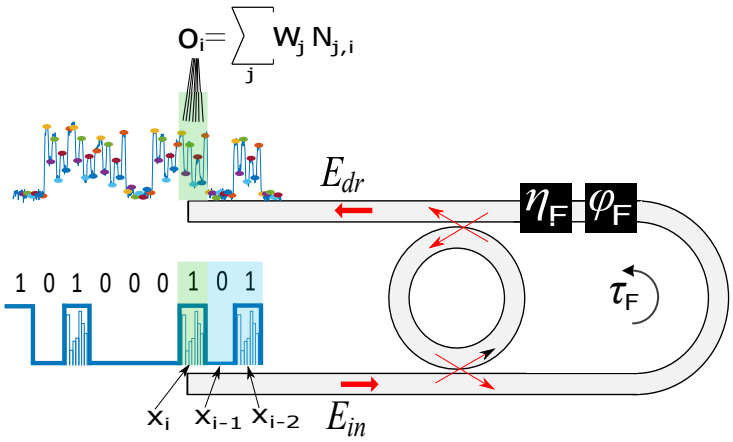

Figure 6.1: Delayed-boolean task scheme. The boolean operation is carried out between the actual bit x_i and a bit x_{i-d} previously input to the microring. The response of the network is estimated from the dynamical response of the microring system, only to the current bit. In this way, the memory properties of the microring resonator are tested.

the output layer of the system, which does not necessarily rely on the memory properties of the real node. From now on, every task will be labeled by the correspondent name, followed by a number $d = 0, 1, 2 \ldots$, indicating that the task involves the current bit x_i and a delayed bit x_{i-d}, already presented to the network.

The performances on each task are computed in terms of the Bit Error Rate (BER), i.e. the number of misclassified bits over the total number of bits processed. To estimate the BER, the output series o_i need firstly to be digitized, and then compared with the digital target y_i. To do so, the output series o_i is normalized. Then, each element is considered to be one or zero, according to a threshold θ_{BER}. Different BER estimations are given by spanning θ_{BER} in the range $[0, 1]$, with steps of 0.01. The BER obtained in a task will refer to the minimum BER achieved when scanning θ_{BER}. The procedure is illustrated in Fig. 6.2(a)(b), in two cases where the BER does not reach or reaches error free operation, respectively. The left column shows a comparison between the target and the (normalized) predicted series, and the right column shows the dependence of the BER on the threshold θ_{BER}. When the system fails the prediction, as in Fig. 6.2(a), there is no threshold θ_{BER} associated with zero BER. On the other

hand, when the system is able to correctly predict the target series, as in Fig. 6.2(b), more θ_{BER} values can lead to zero BER. I define $\theta_{BER} - area$, as the normalized percentage area for zero error operation. A larger $\theta_{BER} - area$ indicates a larger separation between the predicted series values, and hence a more robust prediction result. The quantity $\theta_{BER} - area$ will be used in the following to highlight those system configurations, leading to zero BER, that are more robust.

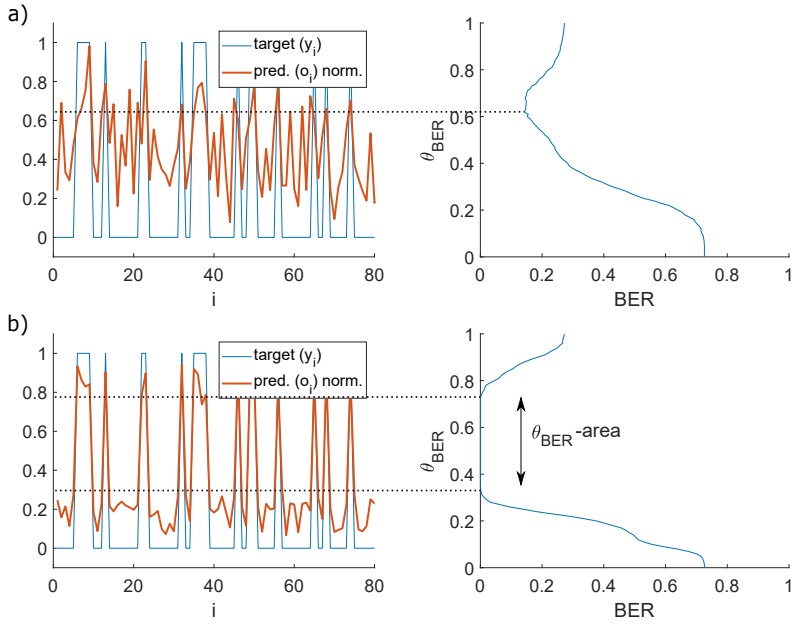

Figure 6.2: Two examples of a) wrong and b) correct task prediction. The left panels show a comparison between the target and normalized predicted series, while the right panels show the dependence of the BER as a function of the threshold θ_{BER}, used to digitize the predicted series.

6.2 Microring resonator with feedback

6.2.1 Setting

The performance on the delayed binary tasks has been evaluated using a processing scheme that differs from the one adopted in the numerical work in section 4.2.4. There, the mask modulation was set at the photon lifetime

($\theta < \tau_{ph}$), and served to couple neighbor virtual nodes using the transient dynamics induced by the microring optical loading time. The mask nodes were applied periodically to every bit processed, whose width, in turn, was stretched to encounter the free carrier lifetime ($b_w \approx \tau_{FC}$). The resulting number of virtual nodes was $N_v = b_w/\theta = 25$, with $b_w = 1ns$ and $\theta = 40ps$. In the experiment, two factors limit the number of actual nodes. First, the limited AWG bandwidth ($500MHz$) did not allow the modulation of the input optical signal at the photon lifetime. Second, the microring has a high free carrier (and thermal) timescale, $\tau_{FC} = 45ns$ (instead of $\tau_{FC} = 3.3ns$, used in the simulations).

The bit width b_w is adapted to the new microring free carrier lifetime, as well as the feedback delay time τ_F. On the other hand, instead of a mask modulation at the photon lifetime, the AWG bandwidth allows to mask the input signal at the free carrier timescale, such to couple virtual nodes using these transients. Nevertheless, the microring is here operated in the linear regime to maximize the linear memory of the system, and this possibility is precluded. Thus, we couple here the virtual nodes by introducing an asynchronous regime. A temporal mismatch between the bit width b_w and the delay time τ_F is applied. In this asynchronized regime, as already explained in section 4.1.2, the feedback allows to connect neighbor virtual nodes that belong to different delay intervals, increasing the complexity of the dynamical response, and in this way the representativity of the network. A temporal mismatch $\tau_F - b_w = k\theta$, with $k = 1$ is applied in all the experiments.

The benchmark tasks have been performed using 3 and 7 virtual nodes, by varying both the delay of the feedback fiber loop (thus its length) and the mask node separation θ. For comparison, the tasks have been also tested with a synchronous regime using 8 virtual nodes. Table 6.2 summarizes all the settings adopted. Similarly to the numerical work, a mask is periodically applied to the input optical signal. A step-wise function constant over intervals of length θ masks every bit, with values that are randomly taken from the set: [0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0]. An example of an optical input signal, representative of the signal that is fed to the microring, is shown in Fig. 6.3. In particular, the full Mach Zehnder modulation is exploited, thus input bits $x_i = 0$ equal an almost zero input optical power (which in the figure is linked to negative voltages from the detector, due to its high electronic gain of 2×10^6 V/W).

From an entire dataset of 4200 elements, the first 200 are used to bring the system on a working regime, while the subsequent 3000 are used for the

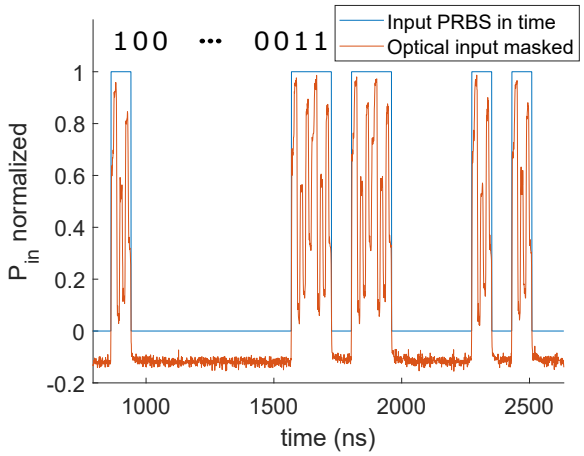

Figure 6.3: Optical input masked before entering the DUT. The mask is composed of seven values and is periodically applied to every bit. This example refers to the setting with $bw = 77ns$ and $\theta = 11ns$.

offline training of the readout weights, and finally, the other 1000 elements serve for testing the network. Note that the input PRBS has order 13 so that the PRBS sequence is periodically repeated every 8192 (2^{13}) elements. In this way, repetitions of the entire sequence over the computing dataset, which has a lower 4200 length, are avoided. Ridge regression is applied at the output layer for optimizing the output weights, with 10^{-4} ridge parameters. Each measure is repeated 3 times, and the correspondent BERs are finally averaged.

6.2.2 SOA nonlinearity

All the tasks are performed using an input power $P_{in} = 78\mu W$, a value well below the appearance of self-pulsing phenomena (refer to Fig. 5.7(b)), that maximizes the linear memory to exploit in the delayed-boolean tasks. Despite this low input optical power, the SOA of the feedback loop can increase the optical power of the circulating signal, introducing nonlinearities into the overall system. To monitor the presence of SOA nonlinear effects, the same task can be repeated with different feedback SOA currents. Then, from the optical response of the system, one can track the state of a virtual node $N_{j,i}$, relative to a particular input element x_i. If the virtual node state varies lin-

θ (ns)	N_v	b_w (ns)	k	τ_F (ns)
19	3	57	1	76
11	7	77	1	88
11	8	88	0	88

Table 6.2: Setting of the input encoding and the feedback delay time adopted for the processing of the benchmark tasks. The feedback delay time is modified by varying the length of the external fiber loop. Input bit duration bw and delay time τ_F are related by $\tau_F - bw = k\theta$, with $k = 0$ indicating the synchronous regime and $k \neq 0$ indicating the asynchronous regime.

early with the SOA current, then the nonlinear behavior is not introduced. To better visualize this, I show in Fig. 6.4 a limited time interval of the microring-based system response, for different feedback SOA currents, reported on top of each panel. Each plot shows the microring response for different detuning values: $\Delta\nu_s = 0$ (in resonance), $\Delta\nu_s = 30GHz$ (out of resonance), $\Delta\nu_s = \pm 6GHz$ (partially in resonance). The masked input optical trace is also shown as a reference, here represented by a logical one, surrounded by zeros. The phase of the feedback signal is in this case not important. Indeed, the optical signal is modulated by exploiting the maximum transmission of the MZI. As a result, the input optical power corresponding to $x_i = 0$ input bits is almost zero, and the resulting drop response is only provided by the feedback signal, thus without interference.

The following observations can be made:

- For $\Delta\nu_s = 0$, the microring-based system response is perfectly aligned with the optical input, and no delayed copies appear.
- For $\Delta\nu_s = 30GHz$, the input signal is out of resonance, and the relative optical trace is delayed by $\tau_F = 88ns$.
- For $\Delta\nu_s = \pm 6GHz$, echoes of the input signal appear.
- Lasing noise starts to appear for SOA currents above $400mA$, with effects already visible at this current for $\Delta\nu_s = 0$.

The state of a virtual node can be sampled and tracked over the different SOA currents and detuning values. As an example, Fig. 6.5 refers to a virtual node sampled at $273ns$, which represents the third virtual node of

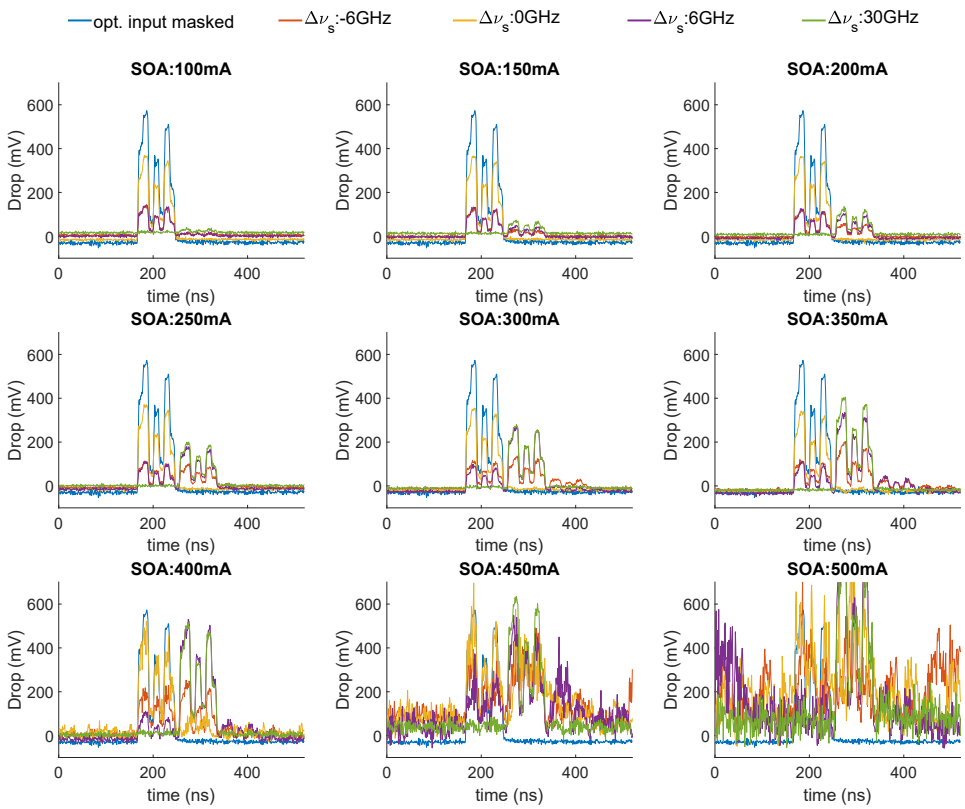

Figure 6.4: Temporal traces obtained when input a PRBS signal to the microring coupled to an external feedback in an asynchronous regime, for varying feedback SOA currents. Each plot shows the input optical signal and the drop response for different detuning values: $\Delta\nu_s = 0$ (in resonance), $\Delta\nu_s = 30GHz$ (out of resonance), $\Delta\nu_s = \pm 6GHz$ (partially in resonance).

the first echo signal, and thus incorporates the effects of the feedback SOA amplification. In the case $\Delta\nu_s = 0$, there is no feedback signal, as the input signal is suddenly dropped. The virtual node state is independent of the SOA current until the system undergoes lasing above $400mA$. For larger detuning values $\Delta\nu_s$, the optical power which is input to the SOA increases, as more optical power at the input port of the microring is transmitted to the through port, and then to the SOA in the feedback loop. For all $\Delta\nu_s \neq 0$, the dependence of the virtual node state is linear with the SOA current, before the lasing effects start to exhibit, and also does not show the effects of SOA gain saturation.

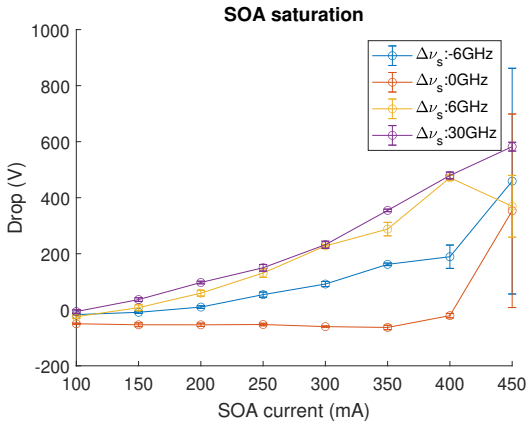

Figure 6.5: Response of a virtual node sampled at $273ns$ as a function of the SOA current, for different starting detuning values.

6.2.3 Computing performance

All the boolean tasks reported in table 6.1 are evaluated from the dynamical response of the microring with external feedback. Several configurations have been explored to find the region of optimal performance. Relying on the previous numerical work, I have explored the influence of the feedback strength, as a function of the electrical current driving the feedback SOA, the feedback phase $\Delta\phi_F$, using the Arduino-based controller, and the starting detuning $\Delta\nu_s$ (here in GHz) between the cold resonance and the input signal frequencies. The SOA current is varied over the full allowed range $[0-500]mA$, the feedback phase varies from almost constructive to destructive interference, i.e. $\Delta\phi_F \in [\pi/6, \pi/3, \pi/2, 2\pi/3, 5\pi/6]$, and $\Delta\nu_s$ ranges in $[-30, +30]GHz$, with $2GHz$ steps. The input optical power is fixed to $P_{in} = 78\mu W$, which, referring to Fig. 5.7, is a value well below the appearance of self-pulsing phenomena.

The results that follow relate to the mask setting with 7 virtual nodes, in an asynchronous regime, where each mask node lasts for a time interval $\theta = 11ns$ and the bit duration is $bw = 77ns$. Since each mask node relates to a correspondent virtual node at the output layer, the microring response is acquired with a picoscope sampling time $T_s = 800ps$. In this way, 14 samples are available per virtual node ($\theta/T_s = 13.75$, is then linear interpolated to obtain 14 samples), from which only the 7^{th} sample indicates the virtual node state. All the other samples are discarded to avoid spurious electronic inter-

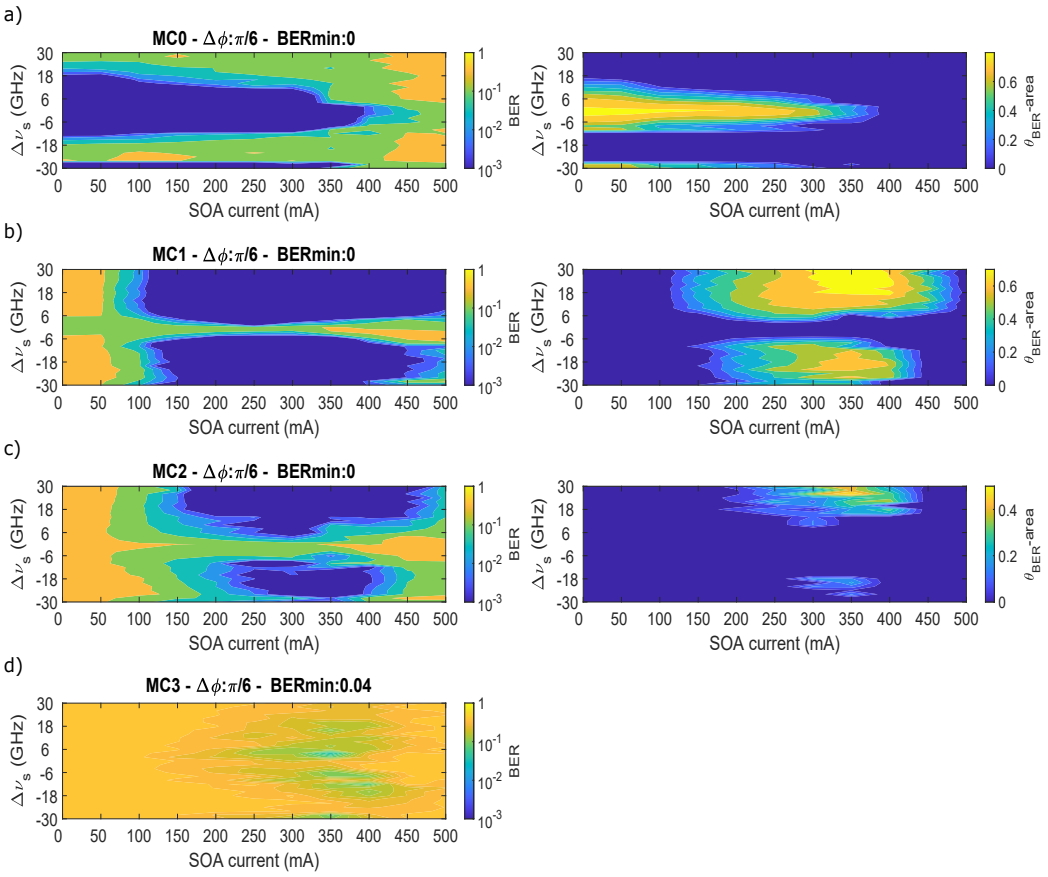

Figure 6.6: BER achieved by the microring-based network in the boolean MC task as a function of the feedback SOA current and the starting detuning $\Delta\nu_s$, at $\Delta\phi_F = \pi/6$. The tasks are the a) MC0, b) MC1, c) MC2, and d) MC3. The left column reports the BER, while the right column reports the threshold area $\theta_{BER} - area$, related to error free operation.

symbol effects, possibly induced by the limited bandwidth of both the input AWG ($500MHz$) and the detector in the output layer task line ($600MHz$).

Before presenting a summary of the best performance obtained in the boolean tasks of table 6.1, some useful examples are reported to better illustrate the physics of the system.

Boolean task MC

The first illustrative task is the boolean memory MC, whose target is the state of the delayed bit. In Fig. 6.6, the results of this task are presented for $d = 0, 1, 2, 3$, and $\Delta\phi_F = \pi/6$. The BER is reported in the left column, while the threshold area θ_{BER}-area, associated with error free operation, is reported in the right column. The x-axis of each plot indicates the feedback SOA current, while the y-axis indicates the starting detuning $\Delta\nu_s$.

In the task MC0, the input bit x_i is also the target. Thus, this task requires the direct transmission of the input to the drop port of the microring, while delayed information is not necessary. Indeed, Fig.6.6(a) indicates that the task is robustly solved with $BER = 0$ and a $\theta_{BER} - area \approx 0.7$, as $\Delta\nu_s$ approaches zero, where the input signal is mostly dropped. The performance start to degrade for SOA currents above $350mA$, when lasing dynamics appear in the system. The task MC1 requires at the drop port also information regarding the bit x_{i-1}, while computing x_i. In this case, $\Delta\nu_s = 0$ is no more a solution, since there is no delayed information at the drop port, as shown in Fig. 6.6(b). Instead, for $\Delta\nu_s \neq 0$ part of the information is transmitted through the feedback loop, and thus delayed at the drop port to better solve the task. Anyway, a minimum feedback SOA current is needed in this case. Still, Fig. 6.6(b), shows that the configurations with the highest $\theta_{BER} - area \sim 0.7$ in the $MC1$ task, are achieved when the input signal frequency is completely out of resonance and the feedback SOA current approaches $400mA$. In this case, indeed, the optical path bypasses the ring, and the drop signal is a $1\tau_F$-delayed version of the input signal, thus matching the requirements of the MC1 task. As a result of the asynchronized regime ($\tau_F - bw = \theta$), this region is also optimal for solving the MC2 task as the first virtual node carries information regarding the response to two previous bits. The MC2 results are shown in Fig. 6.6(c). The task MC3 is not solved with zero BER. The regions of lower BER are here restricted (Fig. 6.6(c)), requiring a pump frequency coupled to the resonance. This last condition, indeed, allows the light signal for multiple iterations along the microring-feedback optical circuit, which is the only way to provide three bits of memory into the system.

While in the previous examples, all the results were referring to a precise feedback phase configuration, the next two examples will focus on the effects of different $\Delta\phi_F$ values.

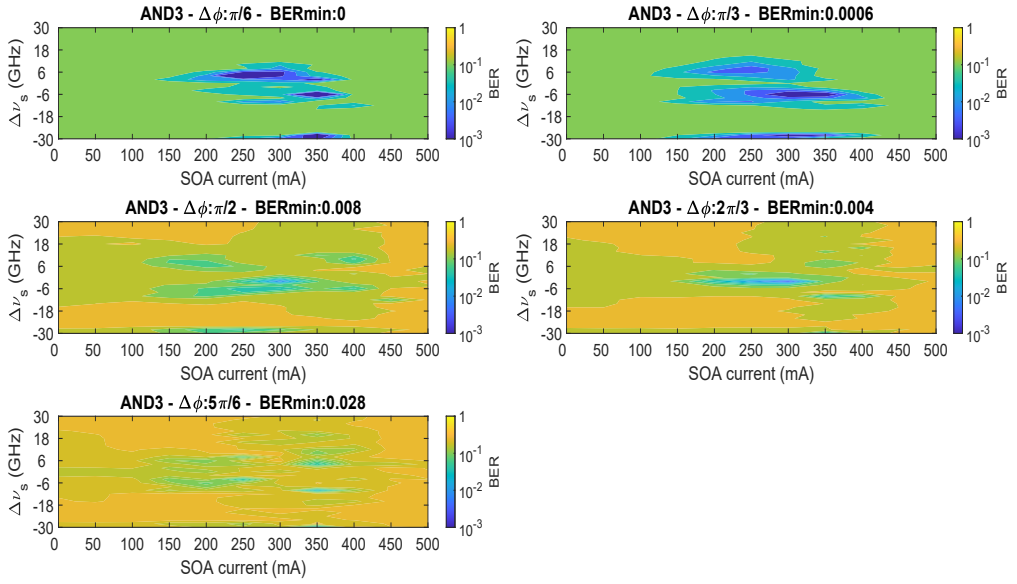

Figure 6.7: BER results for the $AND3$ task, for five values of $\Delta\phi_F$. The minimum BER, reported at the top of each panel, as a constructive interference condition at the drop port, $\Delta\phi_F = \pi/6$, is reached.

Boolean task AND

The results obtained in the boolean AND3 task are reported in Fig. 6.7, for the five values investigated of $\Delta\phi_F$. Each panel reports the BER value as a function of the feedback SOA current and the starting detuning $\Delta\nu_s$. On top of each panel, the minimum BER value obtained for the correspondent $\Delta\phi_F$ configuration is also reported. An attribute of the AND3 task is the dependence of the best BER on the parameter $\Delta\phi_F$, which decreases as the parameter $\Delta\phi_F$ moves from destructive ($\Delta\phi_F = 5\pi/6$) to constructive ($\Delta\phi_F = \pi/6$) interference. Considering that the target value of the AND task is one only when both arguments are one, this fact is reasonable: the system exploits the constructive interference between the drop response of the input bit $x_i = 1$, and the delayed information relative to the input bit $x_{i-3} = 1$ (optical information provided by the feedback loop), to obtain the highest output power. For all the other combinations of x_i and x_{i-3} the output power will be lower, and as a result, the task is solved.

In Fig. 6.8, the results achieved in many delayed AND-d tasks are re-

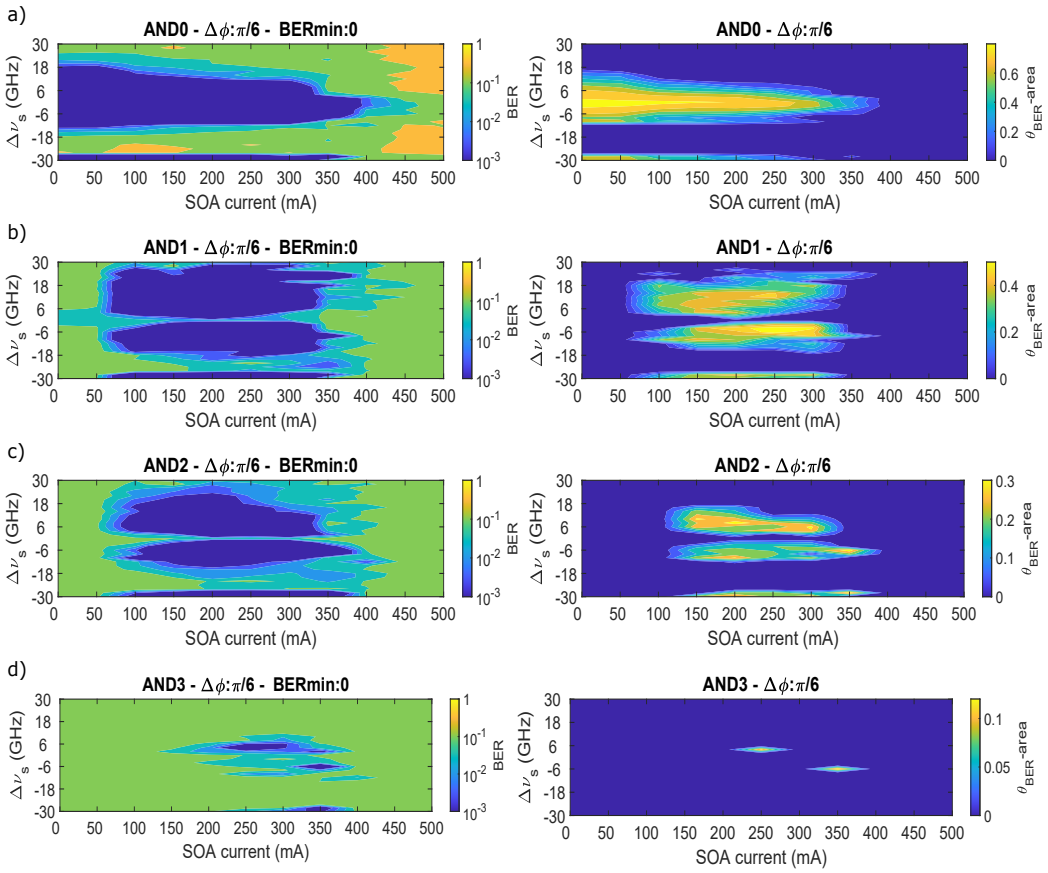

Figure 6.8: BER in different delayed AND-d tasks, with $d = 0, 1, 2, 3$ (left column) and the correspondent $\theta_{BER} - area$, related to error free operation.

ported, with $d = 0, 1, 2, 3$, while considering always the parameter $\Delta\phi_F = \pi/6$. The left column of panels reports the BER, while the right one reports the correspondent $\theta_{BER} - area$, leading to error free operation. The AND0 task requires at the drop port only the information regarding the actual input x_i. For $x_i = 1$ the target is one, for $x_i = 0$ the target is zero. Clearly, the best region in Fig. 6.8(a), where this task is solved, is on resonance, as indicated by the correspondent $\theta_{BER} - area$ at $\Delta\nu_s = 0$. Other configurations that also provide delayed information through the feedback degrade the performance, while configurations providing only delayed information ($|\Delta\nu_s| >> FWHM$) do not solve the task. The AND-1-2 tasks, shown in Fig. 6.8(b)(c) respectively, are solved, instead, even when the pump wave-

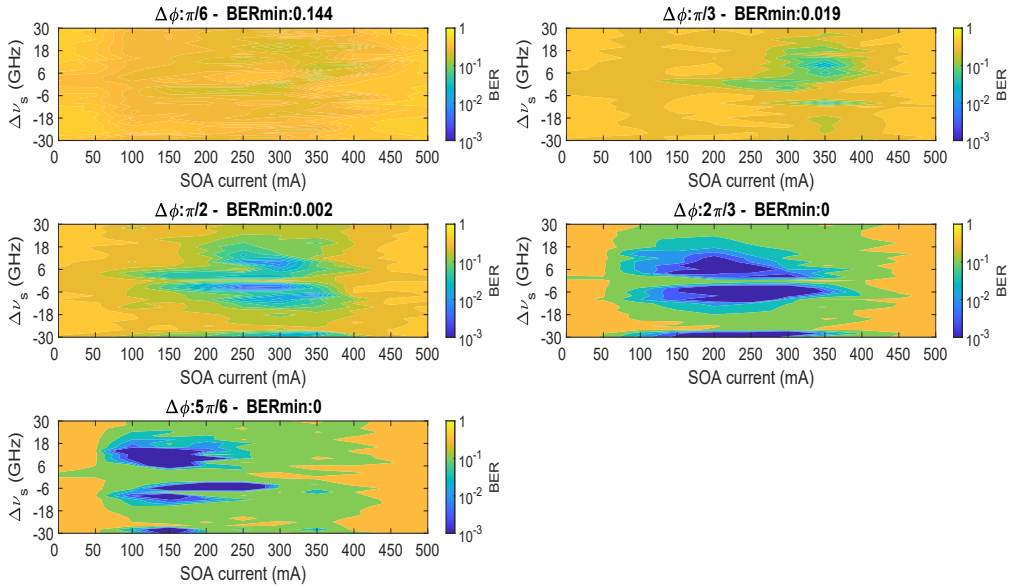

Figure 6.9: BER results for the XOR1 task, for five values of $\Delta\phi_F$. The minimum BER, reported at the top of each panel, decreases as a destructive interference condition at the drop port, $\Delta\phi_F = 5\pi/6$, is reached.

length is out of resonance, relying on the asynchronization between the input and delay optical signals. Nevertheless, more robust configurations emerge towards the region where the pump wavelength is coupled to the microring resonance. Here, the constructive interference occurring at the drop port leads to more robust configurations with error free operation, as indicated by the correspondent $\theta_{BER} - area$ values.

In the case $d = 3$, shown in Fig. 6.8(d), the task is solved in a narrow region, corresponding to the one with the highest number of delayed copies obtained in Fig. 5.10. For $d = 3$, the only way to provide the required past information for solving the task is by multiple iterations of the optical signal through the microring-feedback circuit.

Boolean task XOR

Similarly to the AND task, I report the results for the boolean XOR1 operation in Fig. 6.9, for five values of $\Delta\phi_F$. Each panel reports the BER as a function of the feedback SOA current and the starting detuning $\Delta\nu_s$. By tracking

the value of the best BER, on top of each panel, as a function of $\Delta\phi_F$, an opposite trend with respect to the AND task appears. This time, destructive interference at the drop port is desirable. This is again reasonable, since for the combinations of input $(x_i, x_{i-d}) = (0,0), (1,1)$, having target $y_i = 0$, the correspondent drop optical power can be lower, and thus separable, with respect to the drop optical power of the input combinations $(x_i, x_{i-d}) = (0,1), (1,0)$, having target $y_i = 1$. The results for the XOR-d task, with $d = 1, 2, 3$, are

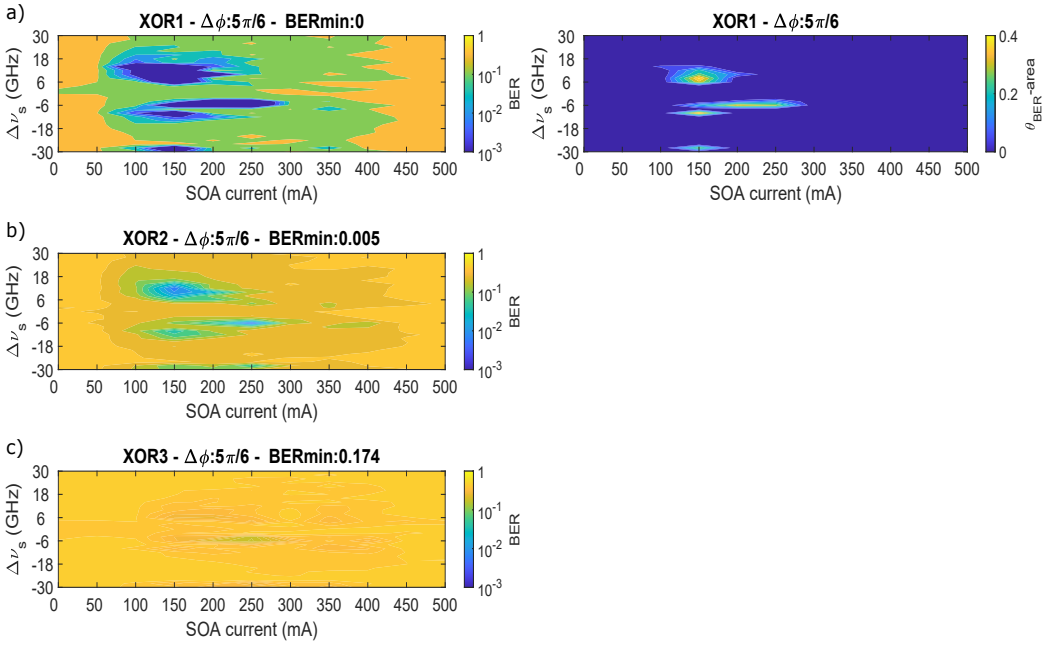

Figure 6.10: BER in different delayed XOR-d tasks, with $d = 1, 2, 3$ (left column) and the correspondent $\theta_{BER} - area$, related to error free operation.

shown in Fig. 6.10(a)(b)(c) respectively, keeping this time $\Delta\phi_F = 5\pi/6$. Differently from the boolean AND and MC operations, the XOR task requires a nonlinear memory to be solved. The simple 1-bit linear memory emerging from the delayed bit at $|\Delta\nu_s >> FWHM|$, or the 2-bit linear memory emerging due to asynchronization, still at $|\Delta\nu_s >> FWHM|$, is here not enough. As a result, for all d-values, the task is never solved when the pump wavelength is out of resonance. Destructive interference at the drop port is here always required, between the input optical information x_i that is suddenly dropped, and the delayed x_{i-d} bit coming from the feedback loop. Hence, for any d value, this task relies on the multiple iterations

that the optical field is able to do along the microring-feedback optical circuit.

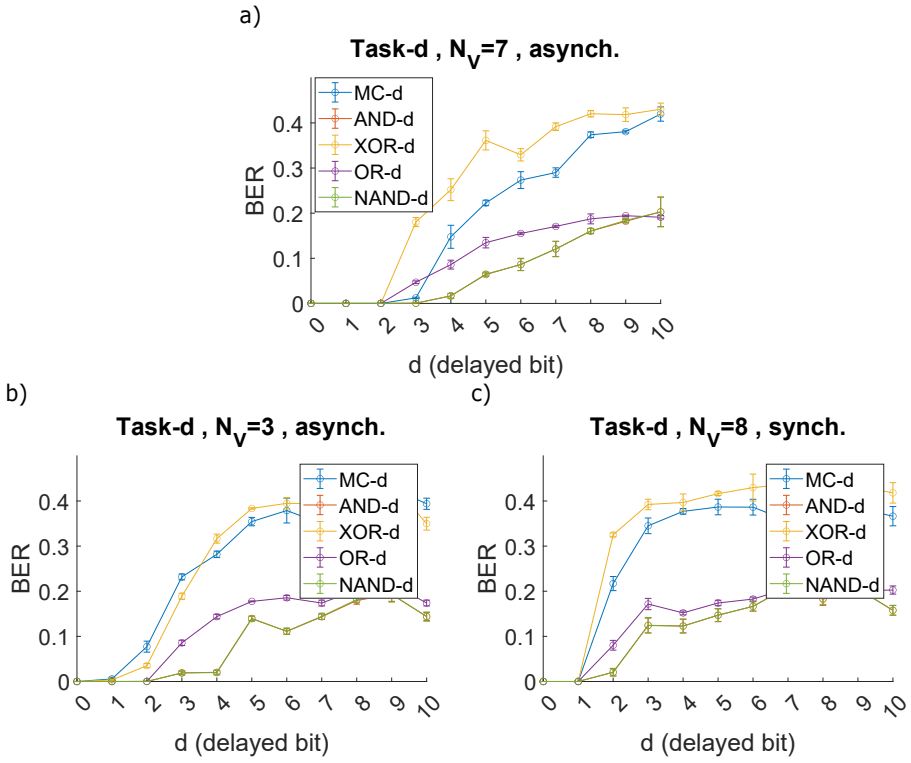

Figure 6.11: Summary of the best BER results obtained on delayed boolean tasks, using the following number of virtual nodes and synchronization regimes: a) $N_V = 3$ and asynchronous regime (k=1), b) $N_V = 3$ and asynchronous regime (k=1), and c) $N_V = 8$ and synchronous regime (k=0).

Results summary

The summary of the lowest BER values for each delay and task performed is reported in Fig. 6.11(a). As can be observed from the boolean MC task, the microring coupled to the feedback loop is able to recall correctly up to three bits in the past. The boolean AND task and its negation NAND task overlap, similarly to what was reported in [129], and can be solved up to $d = 4$, this last with BER=0.02. The OR and XOR tasks are finally solved up to two bits in the past. As already mentioned, these boolean

tasks have been also tested considering both $N_V = 3$ virtual nodes in an asynchronous regime, and $N_V = 8$ virtual nodes in a synchronous regime. Both these measures were done by scanning the detuning in a more restricted area, precisely $\Delta\nu_s \in [-10, +10]GHz$, and are reported in Fig. 6.11(b)(c), respectively. A comparison between the two processing regimes, suggests that operating the system in an asynchronous regime allows better performance on the delayed boolean tasks. In fact, the tasks are only solved up to $d = 1$ for synchronous regimes.

6.3 Microring resonator in absence of feedback

While in the previous section the feedback was a critical component for the correct evaluation of delayed binary tasks, in this section it is disconnected from the microring. The aim is to investigate if the microring resonator alone, thanks to its own nonlinear dynamics, is also able to properly evaluate delayed binary tasks. Recent works have already answered positively to this question. In [130], the free carrier nonlinearity is used to solve the 1-bit delayed OR and AND tasks, but it is not sufficient to solve the 1-bit delayed XOR task. In [131], the authors use two input optical signals, a pump and a probe, coupled to adjacent microring resonances. The microring nonlinearity is triggered by the pump signal, which encodes the input information, and are imprinted to the probe signal, whose samples are given to the ridge regression. The authors show that by using this approach, the microring free carrier nonlinearity can be exploited to solve also 1-bit delayed XOR task. The results in this section show a third possibility to solve the 1-bit delayed XOR task, that exploits the self-pulsation dynamics of the microring, without requiring pump and probe signals. Moreover, these results will serve for a final discussion in section 6.4, where the performance of the microring with and without feedback loop are compared.
The microring that is exploited has radius $r = 7\mu m$, waveguide cross-section of $220 \times 500nm^2$, symmetric gap with the two bus waveguides of $180nm$, $FWHM = 10.5GHz$, a quality factor $Q = 1.8 \times 10^4$, and similar free carrier timescales to the microring used in the feedback experiment (as it belongs to the same chip), estimated as $\tau_{fc} \sim 45ns$.

6.3.1 Setting

The experimental setup for this experiment is the same as reported in Fig. 5.5 without the feedback loop and the phase controller system. The performance of the single microring in the tasks is investigated by spanning the input power P_{in}, the starting detuning $\Delta\lambda_s$, the input bit rate, and the modulation amplitude. Precisely, the parameters span over the following range of values:

- Input power to the microring P_{in}: 0.07, 0.29, 0.51, 0.73, 0.96, 1.18, 1.40, 1.62, 1.84, 2.06 mW.
- $\Delta\lambda_s \in [-20, +20]GHz$ with step of $\sim 1GHz$, for a total of 40 values spanned.
- bit rate: 1, 2.5, 5, 10, 15, 25, 40, 60, 80 Mbps. Correspondingly, the oscilloscope sampling time are set to: 100, 40, 20, 10, 4, 4, 2, 2, 1 ns.
- modulation amplitude: 0.02, 0.04, 0.1, 0.5, 0.7.

The modulation amplitude is set by operating in quadrature the input Mach-Zehnder modulator and only changing the V_{pp} applied on it, through the AWG (see Appendix B). A modulation amplitude of 1 indicates the full use of the Mach-Zehnder transfer function, with zero-bit levels encoded in the minimum power to the Mach Zehnder, and one-bit levels to full transmission. Intermediate amplitude modulations refer to the partial use of the Mach-Zehnder transfer function with respect to the full transmission range, always centered around the quadrature point.

Note that the input bit rate was a parameter not spanned when the feedback was present since it was constrained to the feedback delay time τ_F. Moreover, now there is no mask applied at the input layer, while 5 virtual nodes are sampled at the output layer, synchronously with the input bit injection.

6.3.2 Task results

A summary of the obtained results is reported in Fig. 6.12. Each panel shows the best BER as a function of the input bit rate, for different modulation amplitudes. The results refer to the case $d = 1$, i.e. to tasks operated between the actual bit x_i and the previous bit x_{i-1}. Larger d values are not shown, since they never allow zero BER. Fig. 6.12 highlights the following points:

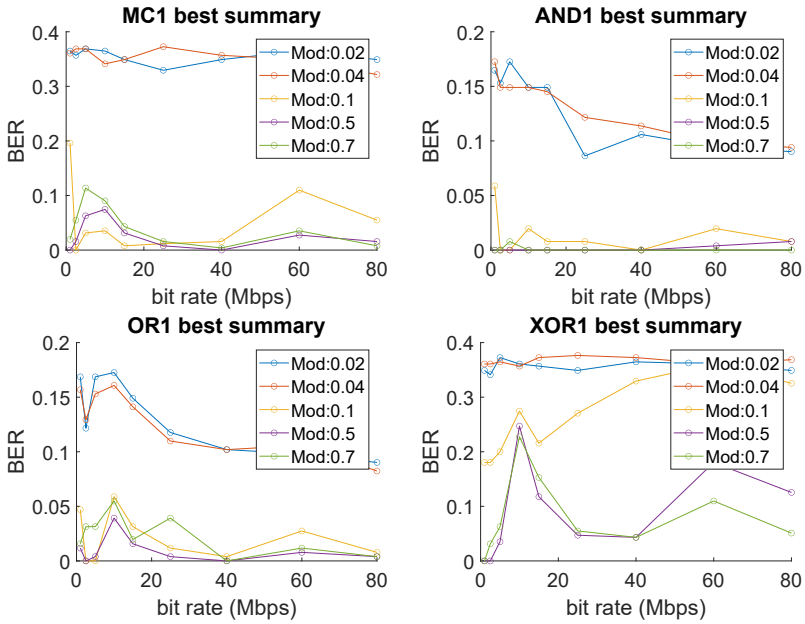

Figure 6.12: Best BER obtained in delayed binary tasks ($d = 1$) using a single microring resonator without feedback loop. The parameters investigated are the starting detuning $\Delta\nu_s$, the input optical power P_{in}, the input bit rate, and the modulation amplitude (Mod).

- The lower modulation amplitudes (0.02, 0.04) do not allow to solve any task, independently of the input bit rate adopted. The optical zero and one levels are close enough to be confused within the readout noise.

- Larger modulation amplitudes (0.1, 0.5, 0.7) allows for zero BER in all tasks. Moreover, the BER curves seem to exhibit a trend, according to which the BER is minimized towards the lowest bit rates (1 Mbps). It then degrades and is later minimized again at bit rates around 40 Mbps. This last bit rate is linked to transient dynamics induced by free carrier nonlinearities, while the slowest bit rates relate to a combination of free carrier and thermal nonlinearities, that are present even in self-pulsing dynamics.

- The AND1 task, appears easier to solve at all bit rates, while the XOR1, which is also the only nonlinear task, is solved only at the slowest input bit rates.

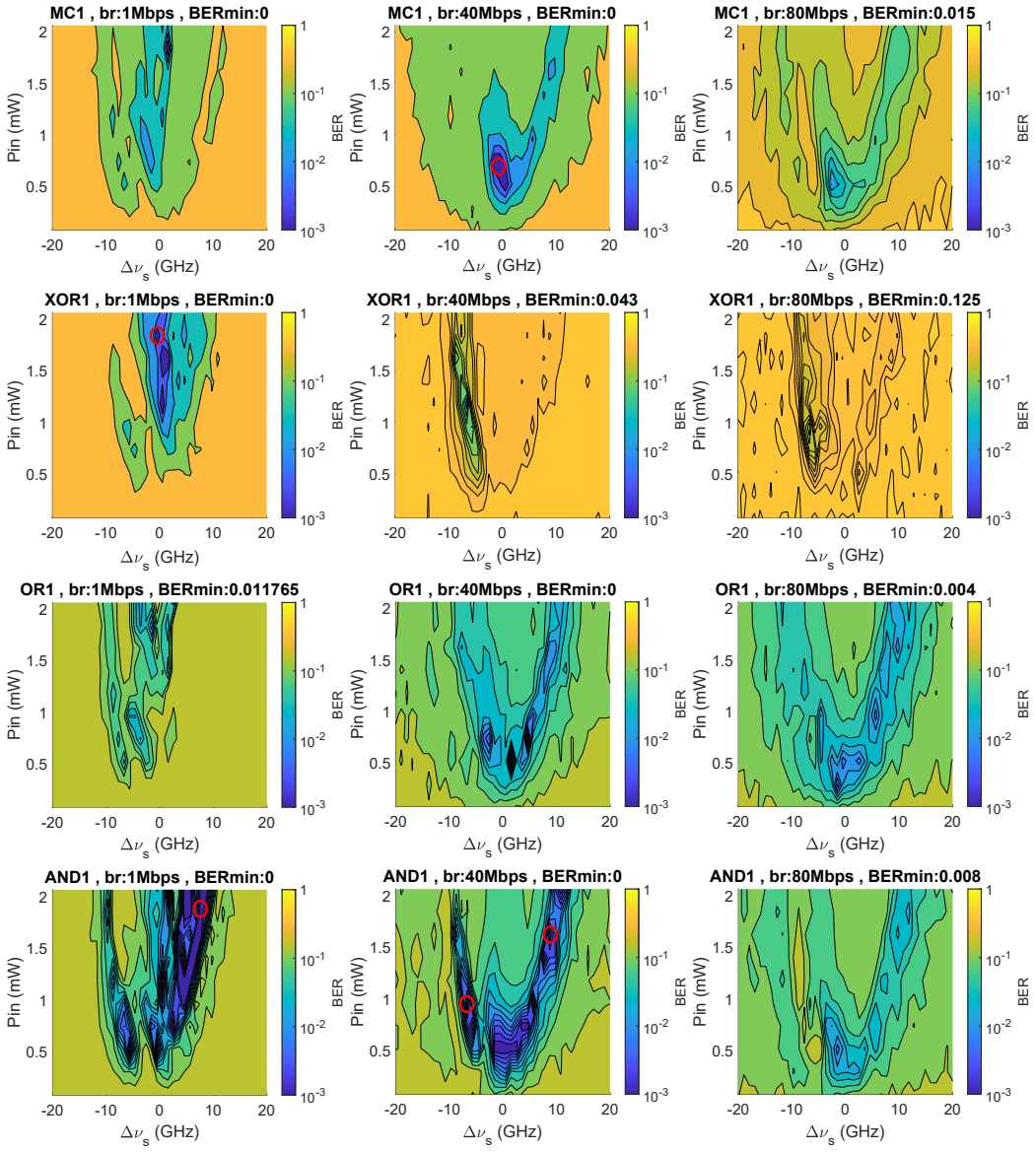

Figure 6.13: BER obtained in delayed binary tasks ($d = 1$) using a single microring resonator without external feedback. Each panel shows the BER performances as a function of the starting detuning $\Delta\lambda_s$ and the input optical power P_{in}. The three columns of panels, from left to right, refer to 1 Mbps, 40 Mbps, and 80 Mbps input bit rates (br), respectively. The amplitude modulation is 0.5. Red dots indicate (P_{in}, $\Delta\nu_s$) configurations with zero BER, whose temporal dynamic is shown in Fig. 6.14.

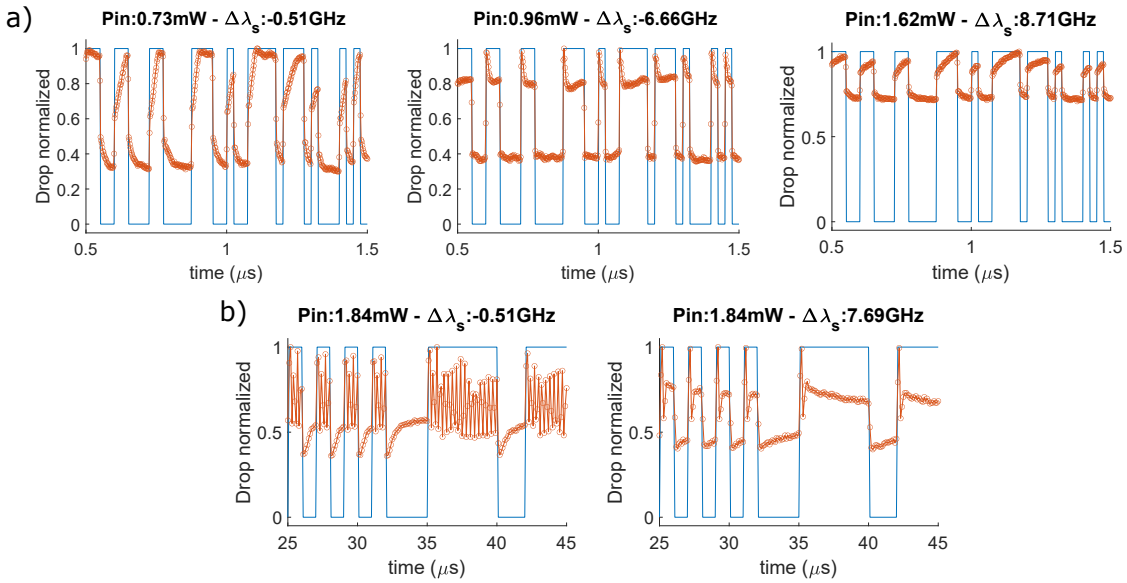

Figure 6.14: Example of experimental traces acquired at an input bit rate of a) $40Mbps$ and b) $1Mbps$, for specific configurations $(P_{in}, \Delta\nu_s)$ indicated on top of each panel. These configurations refer to the red dots in Fig. 6.13. The modulation amplitude is 0.5.

To verify that optimal BERs rely on the microring nonlinearities, the results are shown in a $P_{in} - \Delta\nu_s$ map in Fig. 6.13, for a specific amplitude modulation of 0.5. The three columns relate to three different input bit rates (1, 40, and 80 Mbps, from left to right, respectively). As can be seen, the regions with the best performance are found where the input signal wavelength is coupled to the microring resonance, and improve for higher input average power, task-dependent, in agreement with [130][131].

Some examples of experimental traces are also reported in Fig. 6.14, to highlight the transient dynamics induced by the microring nonlinearities. The traces refer to system configurations indicated with a red circle in Fig. 6.13, leading to error free operation. The traces show clearly the presence of self pulsing, for both input bit rates of 40 Mbps (Fig. 6.14(a)) and 1 Mbps (Fig. 6.14(b)). This last, in particular the trace related to $P_{in} = 1.84mW$ and $\Delta\nu_s = -0.51GHz$, shows that self-pulsation dynamics can be exploited to solve the XOR1 task, thus extending the results obtained in [130][131]. It is interesting to track the behavior of the drop signal in correspondence with a zero-level input bit. The optical signal is still affected by nonlinear-

ities that were previously triggered and that now are relaxing down. It is also interesting to note how the number of spikes in Fig. 6.14(b) depends on the starting detuning $\Delta\lambda_s$, which reduces to one for the positive frequency detuning displayed (section 3.3.2). In particular, Fig. 6.14 suggests a possible strategy for encoding information in the number of microring spikes, similarly to [132] where a numerical cognition task is solved by encoding the input number given to a single integrate-and-fire neuron model, in its spiking nonlinear response.

As a result, the inertia of the thermal and free carrier nonlinearities are here exploited to preserve information about the previous input bit, and consequently for solving the 1-bit delayed binary tasks presented. The finite timescale of the nonlinearities, under this perspective, acts as a memory source for the microring-based reservoir, provided that the input bit rate codifies the information to process at similar timescales to introduce transient dynamics. Dynamical transients generated by only the free carriers or from both the free carriers and thermal effects in self-pulsation dynamics, are useful for computation.

6.4 Discussion

Despite the apparent simplicity of the logical operations considered (AND, OR, XOR, MC, NAND), these tasks usually serve as benchmark tasks because these unveil the available nonlinearity and memory mechanisms of the system. The experiments carried out show that different memory sources are present in the system, with and without the feedback loop, already predicted in the numerical work in chapter 4.

The microring coupled to an external loop relies on the feedback signal, opportunely tuned, for achieving error-free operation in delayed boolean tasks, that require up to 3 bits of memory ($d = 3$). The richness of the system response is here enhanced by masking the input signal and using an asynchronous regime ($\tau_F - bw = \theta$). The AND and XOR tasks previously illustrated, show that the interference condition at the drop port between the actual bit and the delayed bit of interest, is essential in solving the tasks, thus suggesting the importance of the photodetection square low nonlinearity when combined with a feedback loop that provides the optical memory of the system.

On the other hand, when disconnecting the feedback loop, the microring

is still able to achieve error-free operation in delayed boolean tasks, but only considering up to 1 past bit ($d = 1$). In this case, the source of memory relies on the inertia of the thermal and free carrier nonlinearities, properly triggered by input bit rates that introduce transient dynamics, and hence memory, in the drop response.
Thus, memory can emerge within the system both as an optical delayed signal propagating along the feedback loop, and as the state $(\Delta N, \Delta T)$ of the microring resonator, nonlinearly modified by the input optical signal.

Task performance comparison The performed experiments highlight an important difference between the two memory schemes above described. The response of the single microring without feedback loop needs a non-zero optical signal where imprinting the microring nonlinear effects, and thus write the memory information of the system. Thus, logical input bits such as $x_i = 0$ and $x_i = 1$ in a PRBS, need both to be encoded in non-zero optical signals, for probing the past information stored in the microring state. In section 6.3, for example, the modulation amplitude was always < 1, meaning that 0-level input bits were codified in a non-zero optical signal. Otherwise, a pump and probe strategy can be operated [131]. On the other hand, the microring with an external fiber loop is free from this constraint. In fact, even for zero optical input signal, there can be a positive optical power at the drop port due to the delayed optical signal. As a result, the system was operated by exploiting the full Mach Zehnder modulation range at the input layer (modulation amplitude 1) to maximize the signal-to-noise ratio, hence encoding 0-level bits belonging to the input PRBS in almost zero optical power pulses.

Between the investigated tasks, the boolean XOR is considered the hardest delayed binary task. Indeed, in addition to memory requirements, it also needs a proper nonlinear transformation. The results obtained show that both the microring with and without the external delay loop are system configurations that are able to solve the XOR task, although with different processes. The former exploits the feedback parameter $\Delta\phi_F$ to set a destructive interference condition at the drop port of the microring, followed by the photodetection nonlinearity. The latter relies on combinations of thermal and free carrier nonlinearity transients. As a result, the XOR1 is solved in both cases, while the XOR2 is solved only by the microring with an external feedback loop. Referring to the literature, the XOR1 is a nonlinear bench-

mark solved in many other implementations. For example, the XOR1 task is solved in [89] using an integrated reservoir made of 16 microring resonators coupled in a swirl topology, in [129] using an integrated reservoir made only of combiners and splitters, and in [107] with an integrated SOA coupled to a feedback loop operated in a time-delay RC approach. Between the cited works, the XOR2 is solved only in [129] (which extends to even farther delayed bits), thanks to a parallel readout from several nodes in the network. Nevertheless, the XOR2 is a more difficult task when using time-delay RC, and for example, it is not solved in [107]. Thus, solving the XOR2 task with the microring coupled to an external delay loop can be considered a remarkable result.

Task performance improvements The performances obtained in the delayed boolean tasks show that the microring-feedback system has been successfully implemented, and is reliable thanks to a phase-controller system that works in parallel with the apparatus and compensates the environmental noise while performing the task measurements. The system has been investigated with input optical power well below the appearance of self-pulsation, where the microring works in a linear regime, and using an optical feedback delay $\tau_F = 88ns$. This allowed for 7 virtual nodes in the asynchronous regime, and 8 virtual nodes in the synchronous regime. Certainly, the most attractive part of the work comes now, since the system can be upgraded in many ways for solving more complicated benchmark tasks. In the following, I present a list of possible future directions to explore while using this system.

- High input optical power can be investigated for also exploiting the microring nonlinearity. This has proven fundamental in solving the delayed boolean tasks when disconnecting the feedback line, relying both on the free carrier nonlinearity and even self-pulsation dynamics. Thus, using larger input powers also when the microring is coupled to the external loop is desirable, possibly extending the feedback delay up to $1\mu s$, to match the best input rate suggested in Fig. 6.12. Nevertheless, it is worth noting that performing the measures at a high input optical power is not as simple as with the single microring operation, since it also requires the calibration of the phase-controller system. Indeed, both the detector in the slow output line (PMc) and the input ADCs of Arduino, do not have to saturate as a consequence of the increased input power. Attempts to solve the tasks at higher input power have

been already done with the current feedback $\tau_F = 88ns$, but without checking for these saturation problems, and are not discussed in the thesis.

- At larger input power, it is possible to mask the input signal by exploiting the free carrier nonlinearity transients. This last, in combination with an asynchronous regime, may lead to a richer dynamical range in the response of the microring with external feedback, which is beneficial for task computation [133].

- The number of virtual nodes can be increased, by for example extending the length of the fiber loop. In this situation, the virtual nodes can be coupled using a mask signal modulated at the free carrier nonlinearity timescale $\tau_{FC} = 45ns$, thus relying on the microring nonlinearity. A feedback delay $\tau_F = 1\mu s$ and a masking modulation time $\theta = 25ns$ is sufficient in this case to allocate 40 virtual nodes. Alternatively, the microring can be operated in a linear regime and an asynchronous regime can be applied to couple the virtual nodes, as in the experiments performed in this work. This last configuration maximizes the system's linear memory and relies on the photodetection nonlinearity. Note that the realization of the experiment with larger fiber lengths is limited since more environmental phase noise will interest the system.

- Faster phase-controller, employing for example a faster FPGA board (instead of Arduino), voltage amplifier and phase shifter, may allow the stabilization of longer feedback loops (larger τ_F), thus extending the number of virtual nodes available for computation.

- A sufficiently fast AWG, at the input layer, can be operated for masking the input signal at the fastest microring photon lifetime ($\theta < \tau_{ph}$), hence introducing additional transient dynamics. Considering to sample one virtual node per mask node $\theta \sim \tau_{ph}$, as it was proposed in the numerical implementation in section 4.2.3, it is possible to reduce the feedback delay τ_F while maintaining a high number of virtual nodes. In this case, an integration of the feedback loop may become convenient, as it also eliminates the problems related to environmental phase noise encountered during my experiments (section 7.1).

- A higher number of virtual nodes, as suggested by the previous points, and a higher input optical power open the way to compute also analog

tasks, such as the Santa Fe and Mackey-Glass prediction tasks, that were tested in the numerical work. To this aim, a higher input optical power is beneficial also for encoding the analog task information in the optical domain with a higher signal-to-noise ratio.

- More complex structures based on microring resonators, for example CROW and SCISSOR geometries, can be exploited to enhance the real node nonlinearity and thus the richness of the dynamical response.

Chapter 7

Future perspectives

The concepts and strategies learned during this thesis work, have finally led to a series of ideas. Their basic ingredients are presented in the following sections, hoping that they can inspire future applications in the neuromorphic photonic field.

7.1 Hybrid spatial-time delay RC

Realizing a microring resonator with an optical fiber as a delay line brings several complications. As already described in section 5.2, an optical fiber translates environment vibrational, phonic, and thermal noise, into a feedback optical signal characterized by a noisy phase. As a result, a system that compensates for these sources of noise is required and has been implemented in this thesis work. An alternative and attractive way to realize a microring resonator with an external feedback loop is to integrate this last. Within my thesis work, the design of such a structure has been carried out and just fabricated by the foundry AMF (Singapore).

The layout of the design, reported in Fig. 7.1(a), is characterized by a 2cm integrated delay line coupled to a scissor composed of two microrings 7.1(b). Along the delay line, two heaters are included. One of them aims at controlling the phase of the feedback signal, while the second one, included in one arm of an interferometer (section 2.6.4), aims to control the amplitude of the delayed signal. Multimode interferometers MMIs are used to split and combine light where needed. A grating coupler (GC) along the feedback loop also serves to monitor the optical power of the delayed signal. To this

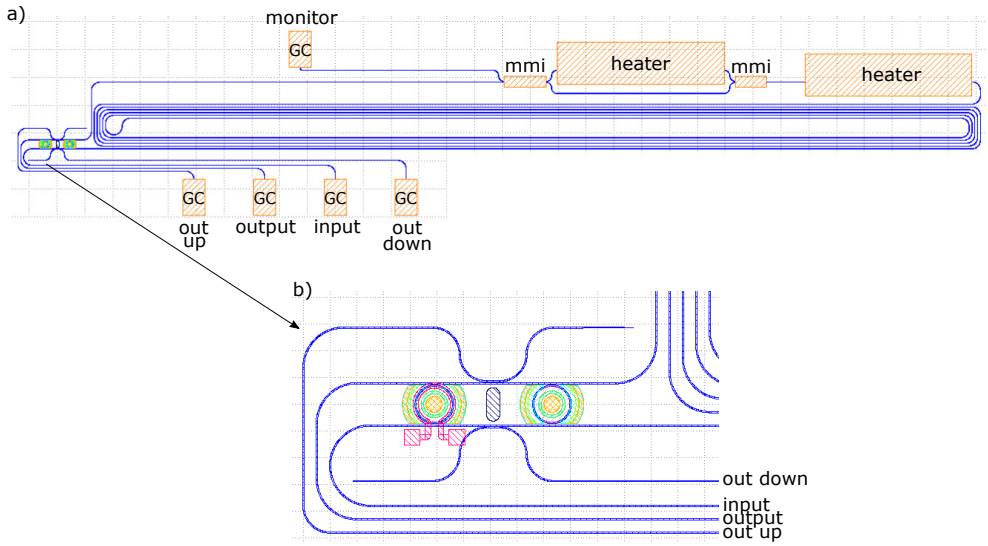

Figure 7.1: a) Layout of the integrated and updated version of the microring resonator coupled to an external feedback loop. b) Focus on the two microrings designed.

purpose, light can be coupled to an external optical fiber and then detected, or an infrared camera can be placed on top of the chip, to directly measure the scattered light by the grating.

The microring is here updated with a SCISSOR (side-coupled integrated spaced sequence of resonator) [134] composed of two identical microrings, having radius $r = 7um$ and a coupling coefficient $k^2 = 0.01$. Both microrings are provided with p-n junctions that aim to control (and in particular reduce) the free carrier timescale. One of the two microrings is also equipped with a heater, to modify the relative resonance positions between the two microrings. In an extreme case, the two microring resonances are totally uncoupled, and the single delayed-microring system is recovered (while using a single wavelength signal).

Finally, it is worth noting that a shorter delay line also allows for a lower number of virtual nodes. Here, the delay $\tau_F \sim 220ps$ can provide around five virtual nodes, using a masking time $\theta = 40ps$ that works at the photon lifetime. In order to improve the number of virtual nodes, the structure is designed with multiple output waveguides, which pick up the signal from different locations of the scissor (drop, out up, out down). In this way, the virtual nodes can be sampled in parallel from several output signals, possibly

increasing their useful number, in a hybrid spatial-time delay RC processing scheme.

7.2 Microring resonator-based fiber sensor

The phase of the feedback signal is a tool for tuning the optical power circulating within the microring. The idea of a fiber sensor coupled to a microring is inspired by Fig. 7.2, where the drop response of the delayed-microring to a CW input signal is acquired as a function of time, without the phase controller, thus at the mercy of the environmental noise. As can be noted, the

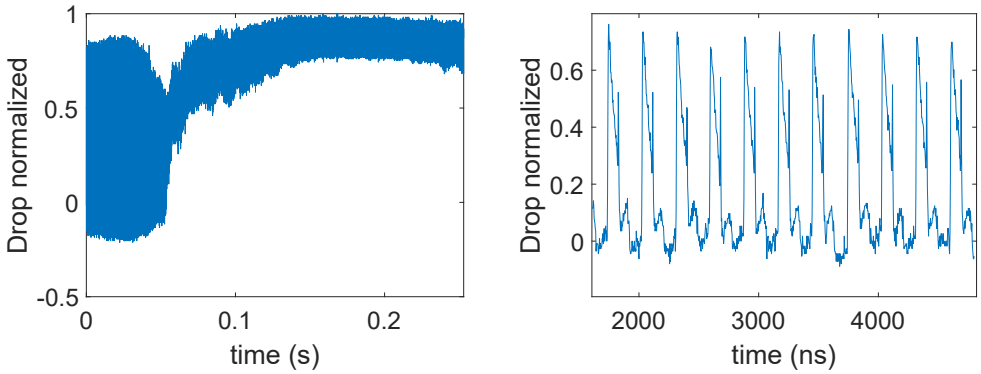

Figure 7.2: a) Drop response to a CW input signal as a function of time, of the delayed microring resonator. b) spiking response in a zoomed time interval. No phase controller is activated, so that the drop signal is subject to environmental phase noise.

dynamic is modified as a function of the time, showing in particular larger oscillation until approximately 0.05s. In other moments, instead, the drop signal is higher and exhibits smaller oscillations. Actually, spiking activities characterized the region with larger oscillations, as shown in a zoom in Fig. 7.2(b). It looks like a behavior induced by the fluctuating phase of the feedback signal, that when induces destructive interference at the drop port (lower signal), it enhances at the same time the optical light circulating across the microring-feedback loop, and thus the microring nonlinearities that lead ultimately to spike emission. This behavior suggests a possible mechanism for using the optical fiber as a sensing element: an agent in the

environment can perturb the fiber, thus varying the phase of the optical signal. The phase perturbation reaches the optical chip where it is translated in a series of spikes emitted by the microring. If the microring is part of a photonic neural network, this process can be used to convert stimulus from the environment into spiking pulses within the neural network. If the phase controller is active, environmental phase noise can be compensated (KHz), and does not participate in the spiking mechanism above described. The target environmental stimulus instead, which should be faster and hence not compensated by the phase controller, induce spiking emission within the integrated microring-based network. The idea extends naturally to other systems that may produce spikes, for example to VCSELs possibly included within 3D networks [135].

7.3 Feed-forward neural network for sensing applications

Different from time-delay RC, but still belonging to the neuromorphic computing field, is the following idea, inspired by a feedforward neural network reported in [79]. The idea is to use a similar scheme for the realization of an integrated neural network for sensing applications. The design concept is schematized in Fig. 7.3. The network exploits the broadcast-and-

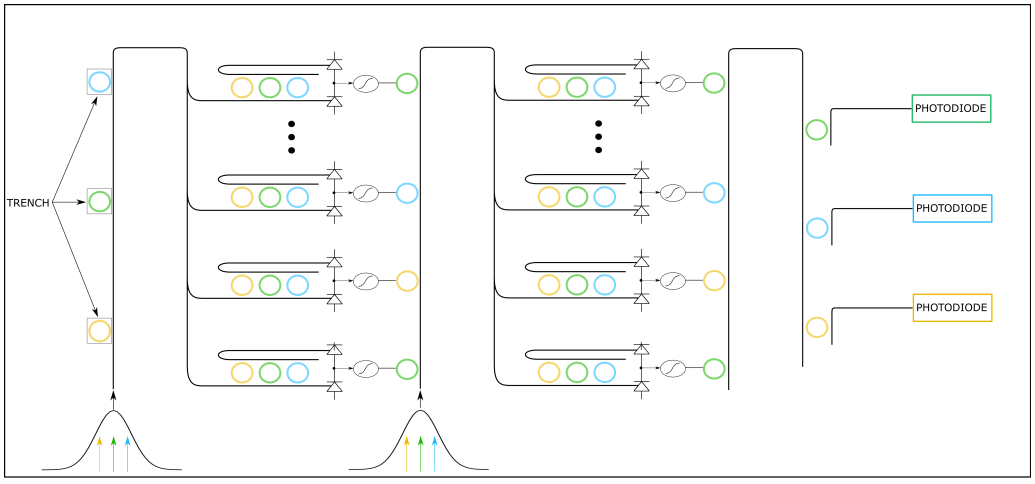

Figure 7.3: Scheme of a prototype FFNN designed for sensing applications.

weight protocol and equips each neuron with a microring weight bank and an electro-optic conversion step that applies the nonlinear transformations using a further microring resonator. Each layer of the network is refreshed by input CW optical signals, either via multiple wavelength channels or via a broadband ASE, which are then modulated by the photonic neurons of the previous layer through their output microrings. The main idea here is to provide the input layer of the FFNN with a set of microring resonators, whose surfaces are engineered to trap specific chemical species. By designing trenches in correspondence of the microrings, for example, these lasts can be exposed to the environment and detect the target chemical species there present. The microring resonance position is shifted in an amount proportional to the concentration of target chemical species, and the information is hence imprinted on the corresponding wavelength channel, thus propagating through the network where it can be processed according to a neural network scheme. Most importantly, the weight banks of each neuron can be trained to perform the task of interest. While in my thesis work, the environment has played the role of a noise source for the delayed-microring network, the same environment is here providing the input information to be processed by the photonic neural network.

Chapter 8

Conclusions

My Ph.D. work focused on the realization of time-delay RC in silicon photonics, in particular relying on the dynamics of silicon microring resonator coupled to an external optical feedback.

During the first phase of the thesis, I became familiar with the linear and nonlinear dynamics that silicon microring resonators exhibit, elaborating a model which also considered a feedback optical signal to the structure. Then, I investigated how to fold a neural network in these dynamics, exploiting all the timescales associated with the microring nonlinearities τ_{fc}, τ_{th} and with optical storage τ_{ph}. The numerical results obtained [95] were a useful tool for understanding reservoir computers containing microring resonators, and constituted the starting points for the design of the subsequent experimental investigation.

During the second phase of the thesis, I set up the optical system for the experimental implementation of the microring resonator with external optical feedback in time-delay RC, obtaining positive results in the delayed boolean benchmark tasks. To this end, it is worth mentioning the effort dedicated to the stabilization of the system against environmental noise. While this last was not accounted for in the numerical work, it appeared a major problem during the experiments, that needed to be solved for ensuring the reproducibility of the results over time. To this aim, I designed and developed a phase-controller system that compensates for the environmental noise, and that can work in parallel with the apparatus while performing task measurements. Moreover, I took part in the design of photonic integrated structures, which include an integrated version of the microring resonator with external optical feedback, and followed the administrative procedures

that lead to their final production. The concept learned during the thesis provided insights for ideas related to future perspective applications, which have been also presented.

Overall, while investigations beyond the single delay node RC architecture are becoming increasingly demanding, these architectures will continue to provide fundamental insights, benefiting from their implementation simplicity. This also stands for the microring resonator with an external feedback loop, where different processing strategies emerged to better solve the benchmark tasks presented in the thesis. However, the experimental investigation presented, which relies on an optical fiber as a feedback loop, could work consistently because stabilized against environmental noise. The high number of virtual nodes that typically time delay RC architectures achieve, by simply extending the feedback loop, is here limited by the environmental noise. Indeed, longer delay times τ_F, and hence longer feedback lengths, for allocating a larger number of virtual nodes, comes at the expense of a larger environmental phase noise along the optical fiber loop, which becomes increasingly difficult to compensate. Instead, the input modulation bandwidth can be increased to enhance the number of virtual nodes. This last, combined with a fully integrated solution of the feedback loop that avoids phase noise, is most probably the best direction to follow for future studies that use a microring resonator with external feedback for time delay RC. This system has been already designed and will be therefore soon investigated.

In conclusion, in this work I presented the role that silicon microring resonators have in photonics neuromorphic applications and extended their functionality also in time delay RC. Microring resonators are certainly an important component for integrated photonic ANN implementations. Indeed microring resonators show a nonlinear response that can serve as a neuron nonlinear activation function, with spiking capabilities. They are basic components in integrated weight banks, which in turn allow the operation and training of the network within time division multiplexing schemes. They can route the information across a network and also act as sensing input nodes for the neural network, if for example their surface is properly engineered. They can also serve as real node for time delay RC. All these ingredients make silicon microring resonators promising candidates within integrated photonic implementation, whose functionalities can work in synergy with other strategic elements, such as VCSEL.

List of publications

Publications

G. Donati, C. R. Mirasso, M. Mancinelli, L. Pavesi, and A. Argyris, “Microring resonators with external optical feedback for time delay reservoir computing,” Optics Express, vol. 30, no. 1, pp. 522–537, 2022.

G. Donati, A. Argyris, C. R. Mirasso, M. Mancinelli, and L. Pavesi, “Noise effects on time delay reservoir computing using silicon microring resonators,” in Integrated Optics: Devices, Materials, and Technologies XXVI, SPIE, vol. 12004, 2022, pp. 219–226.

E. Staffoli, D. Bazzanella, S. Biasi, G. Donati, M. Mancinelli, P. Bettotti, and L. Pavesi, “Nonlinear response of silicon photonics microresonators for reservoir computing neural network,” arXiv preprint arXiv:2302.10203, 2023.

Publications in preparation

S. Biasi, G. Donati, A. Lugnan, M. Mancinelli, E. Staffoli, and L. Pavesi, "Photonic neural networks based on integrated silicon microresonators".

G. Donati, M. Mancinelli, A. Argyris, C. R. Mirasso, and L. Pavesi, "Experimental realization of time delay reservoir computing using a silicon microring resonator with an external optical feedback".

Acknowledgements

I thank Lorenzo Pavesi and Claudio Mirasso for making this cotutelle possible.

I thank Apostolos Argyris for being always available to meet for help and discussing new ideas.

I thank Enya for her incredible support. Her positive attitude and daily calls helped me to proceed in the most difficult moments.

I thank Mirko, for the wonderful hiking adventures in the Trentino mountains.

I thank Susy for the welcome in Majorca the first day I arrived, and then for all the wonderful bicycle explorations we did together. Among the others, we learned an important lesson: independently on how much we can be hungry.. never eat cactus fruits from the streets..

I thank Mattia Mancinelli and Davide Bazzanella for the time spent together in the lab and the remarkable help they provided me.

I thank Stefano Biasi and Riccardo Franchi for the time spent together, especially when designing the latest AMF mask. A sentence here needs to be remembered: "Hello Francois, what about the pdk??".

I thank Alessio Baldazzi for the friendship, that unexpectedly, made us also laboratory neighbors. I also thank Alessio for all the spoilers about Onepiece...

I thank Matteo Sanna, for his incredible humor, making the lab a more funny place where to work. "Arigatou, Mateo San !"

I thank Paolo Bettotti, for introducing me, already during the master thesis, to the neuromorphic photonics field. It was a pleasure to have you as a supervisor in my master's thesis, thanks to your young spirit and energy in discussing new ideas.

I thank all the members of both nanolab and Ifisc for the time spent together.

I thank my family for always being there.

Finally, I thank the Trentino mountains and the sea of Majorca, for the deep peace and the desire to play they transmit.

Funding

The work was partially supported by the Severo Ochoa and Maria de Maeztu Program for Centers and Units of Excellence in R&D (MDM-2017-0711), by the Ministerio de Ciencia e Innovación, and the projects PID2019-111537GB-C21 and PID2019-111537GB-C22. The work was also supported by the European Research Council (ERC) and the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 788793, BACKUP).

Appendix A

Simulation parameters

The corresponding parameters used in the numerical simulations are provided in Table A.1. Most of them refer to [66].

Table A.1: Parameter values used in the numerical simulations in the model of chapter 4 and section 3.3.2.

Parameter	Value	parameter	Value
p	$2\pi \times 6.75\mu m$	λ_o	$1549.66nm$
$1/\tau_o$	$1.68GHz$	$2/\tau_k$	$17.2GHz$
$k^2 = \frac{2\gamma_e p n_g}{c}$	0.01 (n_g=4.1)	$t_r^2 = 1 - k^2$	0.99
a_{rt}	$e^{-\frac{cp\tau_0}{4n_g}}$	Q	3×10^4
dn_{si}/dT	$1.86 \times 10^{-4} K^{-1}$	dn_{si}/dN	$-4.2 \times 10^{-27} m^3$
n_F	1.4682	τ_{ph}	$52.81ps$
τ_{TH}	$83.3ns$	τ_{FC}	$3.3ns$
Γ_c	0.9	V_{eff}	$5.331 \times 10^{-18} m^3$
σ_{FCA}	$1.45 \times 10^{-21} m^2$	η_{FCA}	$\frac{\sigma_{FCA}\Gamma_c c}{2n_{Si}}$
β_{TPA}	$0.79 \times 10^{-11} m/W$	P_{abs}	$\frac{2}{\tau_0}\left\|U(t)\right\|^2$
G_{TPA}	$\frac{c^2\beta_{TPA}}{2V_{eff}n_{Si}^2}$	$FWHM$	$48.4pm$

Appendix B

Optical input encoding

According to the processing scheme adopted in the numerical investigation (section 4.2.3), the input information is encoded in the amplitude of a CW optical signal emitted by an infrared laser, which is then input to the input port of the microring resonator. The encoding in the optical domain is experimentally realized by an electro-optic intensity modulator, specially designed for the transmission of analog signals. The modulator consists of a LiNbO3 MZI (iXblue MXAN-LN-10), having 10 GHz bandwidth. It has one optical input and one optical output, the last providing the modulated waveform desired. The modulator is driven by an RF electronic signal, providing the modulation waveform. It also accepts a secondary electronic bias input, to set the modulator in quadrature. Figure B.1(a) shows the transmission of the device operated in the experiment. Here, the Mach-Zehnder modulator is continuously locked in quadrature by a bias controller (MBC-AN-Board), while the modulation depth is provided by the RF voltage, whose maximum range is achieved at $V = \pm 8$ V.

A pre-compensation of the signal out of the AWG is needed to eliminate the cos^2 dependence of the modulator intensity, especially when analog input series (such as the SantaFe one) are input. To this purpose, the input series undergoes a transformation. First, the input series is normalized with respect to its maximum value, thus resulting in the series $x_i \in [0, 1]$, with $i = 1 \dots N$, being N the series length. Then, a modulation range is chosen by setting the AWG with a certain maximum (V_+) and minimum (V_-) voltage values. To fully exploit this modulation range, x_i needs to be linearly mapped into a new series $u_i \in [-1, 1]$, whose extreme values $u_i = -1(+1)$ will refer to the set voltages $V_-(V_+)$. The problem relies on finding the values u_i, such

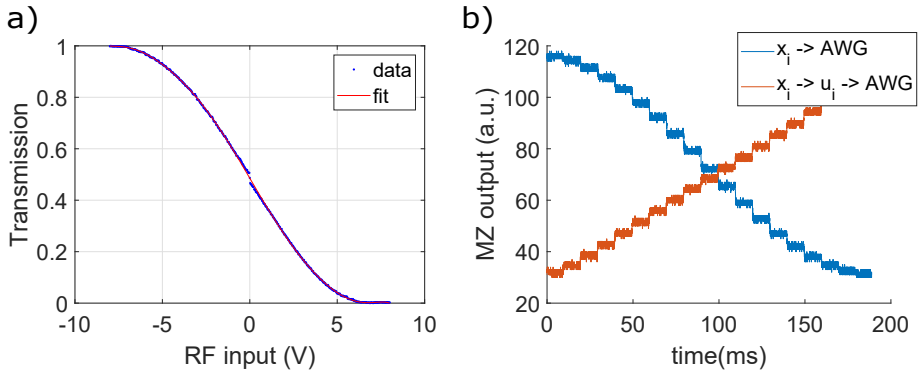

Figure B.1: a) LiNbO3 Mach-Zehnder modulator transmission as a function of the input RF voltage. The modulator is kept in quadrature by a bias controller. b) Linear input series loaded into the AWG with(red curve)/without(blue curve) the pre-compensation needed to eliminate the cos^2 dependence of the modulator intensity.

to linearly transform the input series x_i in the optical domain. Referring to V as the signal out of the AWG, to T as the transmission function of the modulator, and assuming the modulator locked in a negative quadrature as in Fig. B.1, one can impose the following condition:

$$T(V_i) = T(V_+) + \Big(T(V_-) - T(V_+)\Big)x_i, \tag{B.1}$$

where V_i is related to the input series u_i by:

$$V_i = (V_+ - V_-)\frac{(u_i + 1)}{2} + V_-. \tag{B.2}$$

Substituting Eq. B.2 into Eq. B.1,and solving for u_i, leads to:

$$u_i = -1 + \frac{2}{(V_+ - V_-)}\left\{-V_- + T^{-1}\Big[T(V_+) + \Big(T(V_-) - T(V_+)\Big)x_i\Big]\right\}, \tag{B.3}$$

being T^{-1} the inverse function of the modulator transmission. Thus, given the normalized input series x_i and the modulator transmission T, one can estimate the correct series u_i to load into the AWG by using Eq. B.3 (Fig. B.1(b))
In the simpler case $V_+ = -V_-$, and calling $V_{pp} = V_+ - V_-$, Eq. B.3 simplifies in:

$$u_i = \frac{2}{V_{pp}}T^{-1}\Big[T(V_+) + \Big(T(V_-) - T(V_+)\Big)x_i\Big]. \tag{B.4}$$

An alternative approach that avoids the above transformations, is to reduce the V_{pp} range out of the AWG, to the central region of the modulator transfer function, whose behavior is approximately linear. The drawback of this choice is the restricted modulation range used, with respect to the maximum one available.

Appendix C

Arduino-based controller

The physical implementation of a silicon microring resonator coupled to a feedback optical fiber loop makes the system sensible to environmental phase fluctuations. A phase controller has been implemented to compensate for environmental phase noise. The controller exploits an Arduino board to elaborate a correction signal. The board needs to receive the signal to correct v_{in}, to elaborate an error based on a target v_{tar}, and finally elaborate a correction signal v_{out}, similarly to a PID controller. The list of instructions loaded on the Arduino board is presented in code 1. The instructions are written within a void loop, which is continuously repeated, at a rate that defines the correction rate of Arduino. An external analog conversion of v_{out} is then requested by the application. Thus, the correction rate is here optimized by implementing an external DAC with 16 parallel digital input channels, that receive v_{out} from 16 digital output channels present in the Arduino board. The communication via 16 parallel channels boosts the speed, as groups of 8 digital channels (PORTC and PORTC in code 1) can be addressed in parallel, reducing the number of operations per void cycle loop. A DAC predisposed for I2C communication could be also used. In this case v_{out} is communicated as a sequence of 16-bit multiplexed in time, thus requiring 16 Arduino clocks, and leading to lower correction rates.

```
1   void loop()
2   {
3       //setting target from Matlab
4       char matlab_command = Serial.read();
5       Serial.flush();
6       if (matlab_command == 't' )
7       {
8           v_tar = Serial.parseFloat();
9       }
10      // Acquisition signal from the photodetector
11      int v_in = analogRead(A0);
12      // Error estimation
13      err = v_tar - v_pd;
14      // Estimation of the correction signal
15      v_out = (uint16_t) (v_out + (kp*err));
16      // Emission of the correction signal
17      PORTC = v_out;
18      PORTA = (v_out << 8);
19      if (v_out == ByteH)
20      {
21          PORTC = B00111111;
22          PORTA = B11111111;
23      }
24      if (v_out == ByteL)
25      {
26          PORTC = B00111111;
27          PORTA = B11111111;
28      }
29  }
```

Code 1: List of instructions that Arduino continuously repeats over time to keep stable the feedback phase. The frequency at which the void loop is repeated provides also the final correction rate of Arduino, which in this case is approximately $18kHz$.

Acronyms

ADC Analog to Digital Converter. 132

AI Artificial Intelligence. 7

ANN Artificial Neural Network. 7, 8, 11–13, 15, 16, 19–22, 32, 34, 50, 60, 142

ASE Amplified Spontaneous Emission. 98, 139

AWG Arbitrary Waveform Generator. 98, 126, 133, 149, 150

CPU Central Processing Unit. 20

CW Continuous Wave. 98, 149

DAC Digital Analog Converter. 94, 95, 153

DUT Device Under Test. 98, 100

EDFA Erbium Doped Fiber Amplifier. 98

EF Enhancement Factor. 43

FCA Free Carrier Absorption. 36, 37

FCD Free Carrier Dispersion. 36

FFNN Feed Forward Neural Network. 13, 14, 21, 138, 139

FSR Free Spectral Range. 42, 47

FWHM Full Width at Half Maximum. 42, 43

GC Grating Coupler. 100

GPU Grapgic Processing Unit. 7

MMI Multi Mode Interferometer. 135

MZI Mach Zehnder Interferometer. 64, 65, 98, 115, 149

PIC Photonic Integrated Circuit. 22

PID Proportional Integrative Derivative. 93, 153

PRBS Pseudo Random Binary Sequence. 110

RC Reservoir Computing. 1, 8, 16, 18, 21, 53, 59–62, 64–67, 71–74, 76–78, 82, 89, 91, 93, 98, 104, 109, 110, 132, 137, 142

RNN Recurrent Neural Network. 14–16

SL Semiconductor Laser. 66

SNN Spiking Neural Network. 19, 20, 57

SOA Semiconductor Optical Amplifier. 65, 92, 98, 101–103

SOI Silicon On Insulator. 91, 100

SRL Semiconductor Ring Laser. 66

TIR Total Internal Reflection. 23

TLS Tunable Laser Source. 98

TPA Two Photon Absorption. 35–38, 73

TPU Tensor Processing Unit. 7

VCSEL Vertical Cavity Surface Emitting Laser. 65, 138, 142

VOA Variable Optical Attenuator. 98

Bibliography

[1] N. P. Jouppi, C. Young, N. Patil, *et al.*, "In-datacenter performance analysis of a tensor processing unit", in *44th Annual International Symposium on Computer Architecture*, 2017, pp. 1–12.

[2] H. Jaeger, "The "echo state" approach to analysing and training recurrent neural networks-with an erratum note", *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, vol. 148, no. 34, p. 13, 2001.

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

[4] L. Appeltant *et al.*, "Information processing using a single dynamical node as complex system", *Nature communications*, vol. 2, no. 1, p. 468, 2011.

[5] S. Nandakumar, S. R. Kulkarni, A. V. Babu, and B. Rajendran, "Building brain-inspired computing systems: Examining the role of nanoscale devices", *IEEE Nanotechnology Magazine*, vol. 12, no. 3, pp. 19–35, 2018.

[6] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity", *The bulletin of mathematical biophysics*, vol. 5, pp. 115–133, 1943.

[7] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain.", *Psychological review*, vol. 65, no. 6, p. 386, 1958.

[8] M. A. Nielsen, *Neural networks and deep learning*. Determination press San Francisco, CA, USA, 2015, vol. 25.

[9] X. Li and X. Wu, "Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition", in *2015 ieee international conference on acoustics, speech and signal processing (icassp)*, IEEE, 2015, pp. 4520–4524.

[10] H. Sak, A. W. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling", 2014.

[11] J. T. Connor, R. D. Martin, and L. E. Atlas, "Recurrent neural networks and robust time series prediction", *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 240–254, 1994.

[12] W. Maass, "Networks of spiking neurons: The third generation of neural network models", *Neural networks*, vol. 10, no. 9, pp. 1659–1671, 1997.

[13] H. J. Kelley, "Gradient theory of optimal flight paths", *Ars Journal*, vol. 30, no. 10, pp. 947–954, 1960.

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

[15] B. A. Pearlmutter, "Gradient calculations for dynamic recurrent neural networks: A survey", *IEEE Transactions on Neural networks*, vol. 6, no. 5, pp. 1212–1228, 1995.

[16] D. Ha and J. Schmidhuber, "World models", *arXiv preprint arXiv:1803.10122*, 2018.

[17] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey", *Journal of artificial intelligence research*, vol. 4, pp. 237–285, 1996.

[18] C. Fernando and S. Sojakka, "Pattern recognition in a bucket", in *Advances in Artificial Life: 7th European Conference, ECAL 2003, Dortmund, Germany, September 14-17, 2003. Proceedings 7*, Springer, 2003, pp. 588–597.

[19] C. Gallicchio, A. Micheli, and L. Pedrelli, "Deep reservoir computing: A critical experimental analysis", *Neurocomputing*, vol. 268, pp. 87–99, 2017.

[20] E. M. Izhikevich, "Which model to use for cortical spiking neurons?", *IEEE transactions on neural networks*, vol. 15, no. 5, pp. 1063–1070, 2004.

[21] S. Ghosh-Dastidar and H. Adeli, "Spiking neural networks", *International journal of neural systems*, vol. 19, no. 04, pp. 295–308, 2009.

[22] P. A. Bogdan, A. G. Rowley, O. Rhodes, and S. B. Furber, "Structural plasticity on the spinnaker many-core neuromorphic system", *Frontiers in Neuroscience*, vol. 12, p. 434, 2018.

[23] P. A. Merolla *et al.*, "A million spiking-neuron integrated circuit with a scalable communication network and interface", *Science*, vol. 345, no. 6197, pp. 668–673, 2014.

[24] G. E. Moore, "Progress in digital integrated electronics [technical literaiture, copyright 1975 ieee. reprinted, with permission. technical digest. international electron devices meeting, ieee, 1975, pp. 11-13.]", *IEEE Solid-State Circuits Society Newsletter*, vol. 11, no. 3, pp. 36–37, 2006.

[25] M. M. Waldrop, "The chips are down for moore's law", *Nature News*, vol. 530, no. 7589, p. 144, 2016.

[26] P. R. Prucnal and B. J. Shastri, *Neuromorphic photonics*. CRC press, 2017.

[27] X. Lin *et al.*, "All-optical machine learning using diffractive deep neural networks", *Science*, vol. 361, no. 6406, pp. 1004–1008, 2018.

[28] D. Pierangeli, G. Marcucci, and C. Conti, "Photonic extreme learning machine by free-space optical propagation", *Photonics Research*, vol. 9, no. 8, pp. 1446–1454, 2021.

[29] C. M. Valensise, I. Grecco, D. Pierangeli, and C. Conti, "Large-scale photonic natural language processing", *Photonics Research*, vol. 10, no. 12, pp. 2846–2853, 2022.

[30] T. Heuser, M. Pflüger, I. Fischer, J. A. Lott, D. Brunner, and S. Reitzenstein, "Developing a photonic hardware platform for brain-inspired computing based on 5× 5 vcsel arrays", *Journal of Physics: Photonics*, vol. 2, no. 4, p. 044 002, 2020.

[31] Q. Cheng, S. Rumley, M. Bahadori, and K. Bergman, "Photonic switching in high performance datacenters", *Optics express*, vol. 26, no. 12, pp. 16 022–16 043, 2018.

[32] A. L. Washburn, L. C. Gunn, and R. C. Bailey, "Label-free quantitation of a cancer biomarker in complex media using silicon photonic microring resonators", *Analytical chemistry*, vol. 81, no. 22, pp. 9499–9506, 2009.

[33] M. Iqbal *et al.*, "Label-free biosensor arrays based on silicon ring resonators and high-speed optical scanning instrumentation", *IEEE Journal of selected topics in quantum electronics*, vol. 16, no. 3, pp. 654–661, 2010.

[34] A. Z. Subramanian *et al.*, "Silicon and silicon nitride photonic circuits for spectroscopic sensing on-a-chip", *Photonics Research*, vol. 3, no. 5, B47–B59, 2015.

[35] X. Zhang, K. Kwon, J. Henriksson, J. Luo, and M. C. Wu, "A large-scale microelectromechanical-systems-based silicon photonics lidar", *Nature*, vol. 603, no. 7900, pp. 253–258, 2022.

[36] B. E. Saleh and M. C. Teich, *Fundamentals of photonics*. john Wiley & sons, 2019.

[37] K. K. Lee, D. R. Lim, H.-C. Luan, A. Agarwal, J. Foresi, and L. C. Kimerling, "Effect of size and roughness on light transmission in a si/sio 2 waveguide: Experiments and model", *Applied Physics Letters*, vol. 77, no. 11, pp. 1617–1619, 2000.

[38] K. K. Lee, D. R. Lim, L. C. Kimerling, J. Shin, and F. Cerrina, "Fabrication of ultralow-loss si/sio 2 waveguides by roughness reduction", *Optics letters*, vol. 26, no. 23, pp. 1888–1890, 2001.

[39] C. Bellegarde *et al.*, "Improvement of sidewall roughness of submicron soi waveguides by hydrogen plasma and annealing", *IEEE Photonics Technology Letters*, vol. 30, no. 7, pp. 591–594, 2018.

[40] F. Grillot, L. Vivien, S. Laval, D. Pascal, and E. Cassan, "Size influence on the propagation loss induced by sidewall roughness in ultrasmall soi waveguides", *IEEE Photonics Technology Letters*, vol. 16, no. 7, pp. 1661–1663, 2004.

[41] R. Marchetti, C. Lacava, L. Carroll, K. Gradkowski, and P. Minzioni, "Coupling strategies for silicon photonics integrated chips", *Photonics Research*, vol. 7, no. 2, pp. 201–239, 2019.

[42] S. Wu, X. Mu, L. Cheng, S. Mao, and H. Fu, "State-of-the-art and perspectives on silicon waveguide crossings: A review", *Micromachines*, vol. 11, no. 3, p. 326, 2020.

[43] H.-L. Han *et al.*, "High performance ultra-compact soi waveguide crossing", *Optics Express*, vol. 26, no. 20, pp. 25 602–25 610, 2018.

[44] R. Sun *et al.*, "Impedance matching vertical optical waveguide couplers for dense high index contrast circuits", *Optics Express*, vol. 16, no. 16, pp. 11 682–11 690, 2008.

[45] A. V. Tsarev, "Efficient silicon wire waveguide crossing with negligible loss and crosstalk", *Optics Express*, vol. 19, no. 15, pp. 13 732–13 737, 2011.

[46] A. M. Jones, C. T. DeRose, A. L. Lentine, D. C. Trotter, A. L. Starbuck, and R. A. Norwood, "Ultra-low crosstalk, cmos compatible waveguide crossings for densely integrated photonic interconnection networks", *Optics express*, vol. 21, no. 10, pp. 12 002–12 013, 2013.

[47] Y. Zhang *et al.*, "A compact and low loss y-junction for submicron silicon waveguide", *Optics express*, vol. 21, no. 1, pp. 1310–1316, 2013.

[48] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics", *Nature Photonics*, vol. 12, no. 11, pp. 659–670, 2018.

[49] B. Shen, P. Wang, R. Polson, and R. Menon, "An integrated-nanophotonics polarization beamsplitter with 2.4× 2.4 μm2 footprint", *Nature Photonics*, vol. 9, no. 6, pp. 378–382, 2015.

[50] A. Y. Piggott, J. Lu, K. G. Lagoudakis, J. Petykiewicz, T. M. Babinec, and J. Vučković, "Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer", *Nature Photonics*, vol. 9, no. 6, pp. 374–377, 2015.

[51] G. F. Gilestro, G. Tononi, and C. Cirelli, "Widespread changes in synaptic markers as a function of sleep and wakefulness in drosophila", *science*, vol. 324, no. 5923, pp. 109–112, 2009.

[52] G. Cocorullo and I. Rendina, "Thermo-optical modulation at 1.5 mu m in silicon etalon", *Electronics Letters*, vol. 1, no. 28, pp. 83–85, 1992.

[53] M. Mancinelli, D. Bazzanella, P. Bettotti, and L. Pavesi, "A photonic complex perceptron for ultrafast data processing", *Scientific Reports*, vol. 12, no. 1, p. 4216, 2022.

[54] R. Soref and B. Bennett, "Electrooptical effects in silicon", *IEEE journal of quantum electronics*, vol. 23, no. 1, pp. 123–129, 1987.

[55] C. Xiong, D. M. Gill, J. E. Proesel, J. S. Orcutt, W. Haensch, and W. M. Green, "Monolithic 56 gb/s silicon photonic pulse-amplitude modulation transmitter", *Optica*, vol. 3, no. 10, pp. 1060–1065, 2016.

[56] Y. Shen *et al.*, "Deep learning with coherent nanophotonic circuits", *Nature photonics*, vol. 11, no. 7, pp. 441–446, 2017.

[57] A. J. Mercante, P. Yao, S. Shi, G. Schneider, J. Murakowski, and D. W. Prather, "110 ghz cmos compatible thin film linbo3 modulator on silicon", *Optics express*, vol. 24, no. 14, pp. 15 590–15 595, 2016.

[58] V. Sorianello *et al.*, "Graphene–silicon phase modulators with gigahertz bandwidth", *Nature Photonics*, vol. 12, no. 1, pp. 40–44, 2018.

[59] M. Wuttig, H. Bhaskaran, and T. Taubner, "Phase-change materials for non-volatile photonic applications", *Nature photonics*, vol. 11, no. 8, pp. 465–476, 2017.

[60] Z. Fang *et al.*, "Ultra-low-energy programmable non-volatile silicon photonics based on phase-change materials with graphene heaters", *Nature Nanotechnology*, vol. 17, no. 8, pp. 842–848, 2022.

[61] M. Miscuglio, G. C. Adam, D. Kuzum, and V. J. Sorger, "Roadmap on material-function mapping for photonic-electronic hybrid neural networks", *APL Materials*, vol. 7, no. 10, p. 100 903, 2019.

[62] H. Tsang and Y. Liu, "Nonlinear optical properties of silicon waveguides", *Semiconductor Science and Technology*, vol. 23, no. 6, p. 064 007, 2008.

[63] T. Van Vaerenbergh, M. Fiers, J. Dambre, and P. Bienstman, "Simplified description of self-pulsation and excitability by thermal and free-carrier effects in semiconductor microcavities", *Physical Review A*, vol. 86, no. 6, p. 063 808, 2012.

[64] W. H. Pernice, M. Li, and H. X. Tang, "Time-domain measurement of optical transport in silicon micro-ring resonators", *Optics express*, vol. 18, no. 17, pp. 18 438–18 452, 2010.

[65] T. J. Johnson, M. Borselli, and O. Painter, "Self-induced optical modulation of the transmission through a high-q silicon microdisk resonator", *Optics express*, vol. 14, no. 2, pp. 817–831, 2006.

[66] M. Mancinelli, "Linear and non linear coupling effects in sequence of microresonators", Ph.D. dissertation, University of Trento, 2013.

[67] M. Borghi, D. Bazzanella, M. Mancinelli, and L. Pavesi, "On the modeling of thermal and free carrier nonlinearities in silicon-on-insulator microring resonators", *Optics Express*, vol. 29, no. 3, pp. 4363–4377, 2021.

[68] S. F. Preble, Q. Xu, B. S. Schmidt, and M. Lipson, "Ultrafast all-optical modulation on a silicon chip", *Optics letters*, vol. 30, no. 21, pp. 2891–2893, 2005.

[69] V. Van, *Optical microring resonators: theory, techniques, and applications*. CRC Press, 2016.

[70] W. Bogaerts *et al.*, "Optical microring resonators. theory, techniques, and applications", *Laser & Photonics Reviews*, vol. 6, no. 1, pp. 47–73, 2012.

[71] A. Tait *et al.*, "Balanced wdm weight banks for analog optical processing and networking in silicon", in *2015 IEEE Summer Topicals Meeting Series (SUM)*, IEEE, 2015, pp. 110–111.

[72] S. Biasi, P. Guillemé, A. Volpini, G. Fontana, and L. Pavesi, "Time response of a microring resonator to a rectangular pulse in different coupling regimes", *Journal of Lightwave Technology*, vol. 37, no. 19, pp. 5091–5099, 2019.

[73] A. N. Tait *et al.*, "Microring weight banks", *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 22, no. 6, pp. 312–325, 2016.

[74] T. F. de Lima *et al.*, "Design automation of photonic resonator weights", *Nanophotonics*, 2022.

[75] A. N. Tait *et al.*, "Feedback control for microring weight banks", *Optics express*, vol. 26, no. 20, pp. 26 422–26 443, 2018.

[76] C. Huang *et al.*, "Demonstration of scalable microring weight bank control for large-scale photonic integrated circuits", *APL Photonics*, vol. 5, no. 4, p. 040 803, 2020.

[77] W. Zhang *et al.*, "Silicon microring synapses enable photonic deep learning beyond 9-bit precision", *Optica*, vol. 9, no. 5, pp. 579–584, 2022.

[78] A. N. Tait, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Broadcast and weight: An integrated network for scalable photonic spike processing", *Journal of Lightwave Technology*, vol. 32, no. 21, pp. 3427–3439, 2014.

[79] C. Huang *et al.*, "A silicon photonic–electronic neural network for fibre nonlinearity compensation", *Nature Electronics*, vol. 4, no. 11, pp. 837–844, 2021.

[80] A. N. Tait *et al.*, "Neuromorphic photonic networks using silicon photonic weight banks", *Scientific reports*, vol. 7, no. 1, pp. 1–10, 2017.

[81] J. Cheng *et al.*, "A small microring array that performs large complex-valued matrix-vector multiplication", *Frontiers of Optoelectronics*, vol. 15, no. 1, pp. 1–15, 2022.

[82] M. Miscuglio and V. J. Sorger, "Photonic tensor cores for machine learning", *Applied Physics Reviews*, vol. 7, no. 3, p. 031 404, 2020.

[83] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[84] S. Chen, L. Zhang, Y. Fei, and T. Cao, "Bistability and self-pulsation phenomena in silicon microring resonators based on nonlinear optical effects", *Optics Express*, vol. 20, no. 7, pp. 7454–7468, 2012.

[85] G. Priem, P. Dumon, W. Bogaerts, D. Van Thourhout, G. Morthier, and R. Baets, "Optical bistability and pulsating behaviour in silicon-on-insulator ring resonator structures.", *Optics express*, vol. 13, no. 23, pp. 9623–9628, 2005.

[86] V. R. Almeida and M. Lipson, "Optical bistability on a silicon chip", *Optics letters*, vol. 29, no. 20, pp. 2387–2389, 2004.

[87] S. Biasi, R. Franchi, D. Bazzanella, and L. Pavesi, "On the effect of the thermal cross-talk in a photonic feed-forward neural network based on silicon microresonators", *Frontiers in Physics*, vol. 10, p. 1350, 2022.

[88] F. Ashtiani, A. J. Geers, and F. Aflatouni, "An on-chip photonic deep neural network for image classification", *Nature*, pp. 1–6, 2022.

[89] F. Denis-Le Coarer *et al.*, "All-optical reservoir computing on a photonic chip using silicon-based ring resonators", *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 6, pp. 1–8, 2018.

[90] C. Mesaritakis, V. Papataxiarhis, and D. Syvridis, "Micro ring resonators as building blocks for an all-optical high-speed reservoir-computing bit-pattern-recognition system", *JOSA B*, vol. 30, no. 11, pp. 3048–3055, 2013.

[91] C. Mesaritakis, A. Bogris, A. Kapsalis, and D. Syvridis, "High-speed all-optical pattern recognition of dispersive fourier images through a photonic reservoir computing subsystem", *Optics letters*, vol. 40, no. 14, pp. 3416–3419, 2015.

[92] J. Xiang, A. Torchy, X. Guo, and Y. Su, "All-optical spiking neuron based on passive microresonator", *Journal of Lightwave Technology*, vol. 38, no. 15, pp. 4019–4029, 2020.

[93] E. M. Izhikevich, *Dynamical systems in neuroscience*. MIT press, 2007.

[94] T. Van Vaerenbergh *et al.*, "Cascadable excitability in microrings", *Optics express*, vol. 20, no. 18, pp. 20 292–20 308, 2012.

[95] G. Donati, C. R. Mirasso, M. Mancinelli, L. Pavesi, and A. Argyris, "Microring resonators with external optical feedback for time delay reservoir computing", *Optics Express*, vol. 30, no. 1, pp. 522–537, 2022.

[96] D. Brunner, B. Penkovsky, B. A. Marquez, M. Jacquot, I. Fischer, and L. Larger, "Tutorial: Photonic neural networks in delay systems", *Journal of Applied Physics*, vol. 124, no. 15, p. 152 004, 2018.

[97] Y. Paquot *et al.*, "Optoelectronic reservoir computing", *Scientific reports*, vol. 2, no. 1, p. 287, 2012.

[98] A. Argyris, "Photonic neuromorphic technologies in optical communications", *Nanophotonics*, vol. 11, no. 5, pp. 897–916, 2022.

[99] L. Larger *et al.*, "Photonic information processing beyond turing: An optoelectronic implementation of reservoir computing", *Optics express*, vol. 20, no. 3, pp. 3241–3249, 2012.

[100] F. Duport, B. Schneider, A. Smerieri, M. Haelterman, and S. Massar, "All-optical reservoir computing", *Optics express*, vol. 20, no. 20, pp. 22 783–22 795, 2012.

[101] R. M. Nguimdo, G. Verschaffelt, J. Danckaert, and G. Van der Sande, "Simultaneous computation of two independent tasks using reservoir computing based on a single photonic nonlinear node with optical feedback", *IEEE transactions on neural networks and learning systems*, vol. 26, no. 12, pp. 3301–3307, 2015.

[102] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, "Parallel photonic information processing at gigabyte per second data rates using transient states", *Nature communications*, vol. 4, no. 1, p. 1364, 2013.

[103] A. Dejonckheere *et al.*, "All-optical reservoir computer based on saturation of absorption", *Optics express*, vol. 22, no. 9, pp. 10 868–10 881, 2014.

[104] Q. Vinckier *et al.*, "High-performance photonic reservoir computer based on a coherently driven passive cavity", *Optica*, vol. 2, no. 5, pp. 438–446, 2015.

[105] J. Bueno, J. Robertson, M. Hejda, and A. Hurtado, "Comprehensive performance analysis of a vcsel-based photonic reservoir computer", *IEEE Photonics Technology Letters*, vol. 33, no. 16, pp. 920–923, 2021.

[106] K. Takano *et al.*, "Compact reservoir computing with a photonic integrated circuit", *Optics express*, vol. 26, no. 22, pp. 29 424–29 439, 2018.

[107] K. Harkhoe, G. Verschaffelt, A. Katumba, P. Bienstman, and G. Van der Sande, "Demonstrating delay-based reservoir computing using a compact photonic integrated chip", *Optics express*, vol. 28, no. 3, pp. 3086–3096, 2020.

[108] S. Ortin *et al.*, "A unified framework for reservoir computing and extreme learning machines based on a single time-delayed neuron", *Scientific reports*, vol. 5, no. 1, p. 14 945, 2015.

[109] P. Antonik, F. Duport, M. Hermans, A. Smerieri, M. Haelterman, and S. Massar, "Online training of an opto-electronic reservoir computer applied to real-time channel equalization", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 11, pp. 2686–2698, 2016.

[110] F. Stelzer, A. Röhm, R. Vicente, I. Fischer, and S. Yanchuk, "Deep neural networks using a single neuron: Folded-in-time architecture using feedback-modulated delay loops", *Nature communications*, vol. 12, no. 1, p. 5164, 2021.

[111] F. Duport, A. Smerieri, A. Akrout, M. Haelterman, and S. Massar, "Fully analogue photonic reservoir computer", *Scientific reports*, vol. 6, no. 1, p. 22 381, 2016.

[112] A. Argyris, J. Bueno, and I. Fischer, "Photonic machine learning implementation for signal recovery in optical communications", *Scientific reports*, vol. 8, no. 1, p. 8487, 2018.

[113] A. Argyris, J. Bueno, and I. Fischer, "Pam-4 transmission at 1550 nm using photonic reservoir computing post-processing", *IEEE Access*, vol. 7, pp. 37 017–37 025, 2019.

[114] I. Estébanez, S. Li, J. Schwind, I. Fischer, S. Pachnicke, and A. Argyris, "56 gbaud pam-4 100 km transmission system with photonic processing schemes", *Journal of Lightwave Technology*, vol. 40, no. 1, pp. 55–62, 2021.

[115] I. Estébanez, J. Schwind, I. Fischer, and A. Argyris, "Accelerating photonic computing by bandwidth enhancement of a time-delay reservoir", *Nanophotonics*, vol. 9, no. 13, pp. 4163–4171, 2020.

[116] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems", *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.

[117] A. F. Atiya and A. G. Parlos, "New results on recurrent network training: Unifying the algorithms and accelerating convergence", *IEEE transactions on neural networks*, vol. 11, no. 3, pp. 697–709, 2000.

[118] H. Jaeger and H. Haas, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication", *science*, vol. 304, no. 5667, pp. 78–80, 2004.

[119] A. S. Weigend and N. A. Gershenfeld, "Results of the time series prediction competition at the santa fe institute", in *IEEE international conference on neural networks*, IEEE, 1993, pp. 1786–1793.

[120] M. Inubushi and K. Yoshimura, "Reservoir computing beyond memory-nonlinearity trade-off", *Scientific reports*, vol. 7, no. 1, p. 10 199, 2017.

[121] H. Jaeger, "Tutorial on training recurrent neural networks, covering bppt, rtrl, ekf and the" echo state network" approach", 2002.

[122] H. Jaeger, "Short term memory in echo state networks", 2001.

[123] M. C. Mackey and L. Glass, "Oscillation and chaos in physiological control systems", *Science*, vol. 197, no. 4300, pp. 287–289, 1977.

[124] J. B. Moragues, "Photonic information processing", Ph.D. dissertation, University of the Balearic Islands, 2018.

[125] I. Estébanez, I. Fischer, and M. C. Soriano, "Constructive role of noise for high-quality replication of chaotic attractor dynamics using a hardware-based reservoir computer", *Physical Review Applied*, vol. 12, no. 3, p. 034 058, 2019.

[126] M. C. Soriano *et al.*, "Optoelectronic reservoir computing: Tackling noise-induced performance degradation", *Optics express*, vol. 21, no. 1, pp. 12–20, 2013.

[127] G. Donati, A. Argyris, C. R. Mirasso, M. Mancinelli, and L. Pavesi, "Noise effects on time delay reservoir computing using silicon microring resonators", in *Integrated Optics: Devices, Materials, and Technologies XXVI*, SPIE, vol. 12004, 2022, pp. 219–226.

[128] M. C. Soriano *et al.*, "Delay-based reservoir computing: Noise effects in a combined analog and digital implementation", *IEEE transactions on neural networks and learning systems*, vol. 26, no. 2, pp. 388–393, 2014.

[129] K. Vandoorne *et al.*, "Experimental demonstration of reservoir computing on a silicon photonics chip", *Nature communications*, vol. 5, no. 1, p. 3541, 2014.

[130] D. Bazzanella, S. Biasi, M. Mancinelli, and L. Pavesi, "A microring as a reservoir computing node: Memory/nonlinear tasks and effect of input non-ideality", *Journal of Lightwave Technology*, vol. 40, no. 17, pp. 5917–5926, 2022.

[131] M. Borghi, S. Biasi, and L. Pavesi, "Reservoir computing based on a silicon microring and time multiplexing for binary and analog operations", *Scientific Reports*, vol. 11, no. 1, p. 15 642, 2021.

[132] H. Rapp, M. P. Nawrot, and M. Stern, "Numerical cognition based on precise counting with a single spiking neuron", *IScience*, vol. 23, no. 2, p. 100 852, 2020.

[133] T. Hülser, F. Köster, L. Jaurigue, and K. Lüdge, "Role of delay-times in delay-based photonic reservoir computing", *Optical Materials Express*, vol. 12, no. 3, pp. 1214–1231, 2022.

[134] J. E. Heebner, R. W. Boyd, and Q.-H. Park, "Scissor solitons and other novel propagation effects in microresonator-modified waveguides", *JOSA B*, vol. 19, no. 4, pp. 722–731, 2002.

[135] A. Skalli *et al.*, "Photonic neuromorphic computing using vertical cavity semiconductor lasers", *Optical Materials Express*, vol. 12, no. 6, pp. 2395–2414, 2022.