Direct coupling of nonlinear integrated cavities for all-optical reservoir computing

Ivan Boikov[1,*], Daniel Brunner[2] and Alfredo De Rossi[1]

[1]Thales Research & Technology, Palaiseau Cedex, 91767, France

[2]FEMTO-ST Institute / Optics Department, CNRS & University Bourgogne Franche-Comté, Besançon Cedex, 25030, France

20 July 2023

Abstract

We consider theoretically a network of directly coupled optical microcavities to implement a space-multiplexed optical neural network in an integrated nanophotonic circuit. Nonlinear photonic network integrations based on direct coupling ensures a highly dense integration, reducing the chip footprint by several orders of magnitude compared to other implementations. Different nonlinear effects inherent to such microcavities are studied when used for realizing an all-optical autonomous computing substrate, here based on the reservoir computing concept. We provide an in-depth analysis of the impact of basic microcavity parameters on computational metrics of the system, namely, the dimensionality and the consistency. Importantly, we find that differences between frequencies and bandwidths of supermodes formed by the direct coupling is the determining factor of the reservoir's dimensionality and its scalability. The network's dimensionality can be improved with frequency-shifting nonlinear effects such as the Kerr effect, while two-photon absorption has an opposite effect. Finally, we demonstrate in simulation that the proposed reservoir is capable of solving the Mackey-Glass prediction and the optical signal recovery tasks at GHz timescale.

1 Introduction

Artificial Neural Networks (ANNs) are a computing approach different from the conventional algorithmic concepts. They are inspired by the most basic principles of human brains, where numerous neurons are connected with synapses to

---

*Electronic mail: ivan.boikov@thalesgroup.com

create a network. Currently, ANNs are mostly implemented on the von Neumann architecture, i.e. on digital computers, and, while successful in many fields, this approach has its own drawbacks, mainly a data bottleneck inherent to the architecture as well as significant power requirements. In addition, the latency between an input and a computed output presents a major concern in some applications. One example is optical communications, where an exponentially growing demand for internet connectivity requires increasingly faster processing of optical signals [1]. A conventional approach is based on the processing with digital algorithms [2, 3, 4] or more recent ANN-inspired solutions [5]. However, at higher bandwidths, an efficient processing becomes complicated. The development of ANNs in a physical layer can help overcome this challenge [6, 7]. Besides analog electronics [8], the manipulation of optical signals is highly relevant when a large bandwidth operation is important, as was demonstrated in the field of microwave photonics [9]. For implementing such computing systems, integrated photonics is a promising platform that provides compactness and intrinsic nonlinear effects necessary for efficient and scalable ANN in hardware [10, 11].

Reservoir Computing (RC) has attracted a lot of interest in the in-hardware computing field as it provides an implementation-friendly ANN topology with competitive performance and a simplistic training procedure [12]. Reservoir Computing on integrated photonic platforms has been demonstrated in simulations [13, 14] and experimentally [15, 16]. These works considered optical microcavities or splitters as neurons with long intra-neural connections, i.e. with a non-negligible coupling delay. This was required to match the system's internal transient timescales to available electronics [15]. However, these delay lines are costly in terms of a chip footprint and optical losses and limit scalability. As an example, a 280 ps delay requires 2 cm of silicon waveguide [15], which results in approximately one neuron per square millimeter of the chip surface.

In this study, we consider evanescent, i.e. direct, coupling between adjacent cavities as an alternative to create maximally compact integrated all-optical computing systems. We propose a RC consisting of a rectangular grid of directly coupled nonlinear microcavities and study the impact of cavity parameters and reservoir geometry on its basic computational properties, namely the dimensionality [17, 18] and the consistency [19]. We found that a longer photon lifetime allows for a better scalability due to more variability between coupling-induced supermodes of the system. Importantly, we found that the nonlinear effects can impact computing performance in opposite directions, where two-photon absorption (TPA) reduces the reservoir's dimensionality, while free carrier dispersion (FCD) and the optical Kerr effect increase the dimensionality. Moreover, unlike TPA, a sufficiently strong FCD and Kerr effect do induce chaotic dynamics in this purely dissipative network without gain. Finally, we evaluate the RC's performance in Mackey-Glass prediction and optical signal recovery tasks. For simulations we used cavity and waveguide parameters accepted as standard for GaAs integrated photonic circuits. We found that such integrated photonic reservoirs show a comparable performance to other implementations found in the literature, which shows a great potential of this maximum integration-density

2

approach.

2 System of coupled nonlinear microcavities

The backbone of our photonic RC is a $N_\parallel \times N_\perp$ square grid of directly coupled nonlinear cavities (see Figure 1(a)). One side of the grid is coupled to the input waveguide according to cavity-waveguide coupling coefficient κ. Through this waveguide an optical input $u(t)$ is injected. Assuming intra-cavity and cavity-waveguide coupling strengths are weak, the coupled-mode theory can be applied [20]. For a purely linear case, the electric field of microcavity modes $\boldsymbol{a}(t)$ can be described with

$$\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}t} = \left(\mathrm{i}\boldsymbol{\omega} - \frac{\boldsymbol{\Gamma}^{\mathrm{o}} + \boldsymbol{\Gamma}^{\mathrm{e}}}{2}\right) \odot \boldsymbol{a} + (\hat{M}^\mu + \hat{M}^\kappa)\boldsymbol{a} + \boldsymbol{K}^{\mathrm{i}}s(t), \qquad (1)$$

where $\boldsymbol{\omega}$ are the cavities' resonant frequencies, $\boldsymbol{\Gamma}^{\mathrm{o}}$ are intrinsic optical losses, $\boldsymbol{\Gamma}^{\mathrm{e}}$ are optical losses to waveguide, $\odot$ is the the Hadamard product, $\hat{M}^\mu$, $\hat{M}^\kappa$ are direct and waveguide-assisted cavity-cavity coupling matrices, $\boldsymbol{K}^{\mathrm{i}}$ is an input waveguide-cavity coupling matrix (see A). Here, $\hat{M}^\mu_{km} = -(\hat{M}^\mu_{mk})^* = |\mu_{km}|\exp(\mathrm{i}\varphi_{km})$ [21] (Section 7.5) if the k-th and m-th cavities are directly coupled and $K^{\mathrm{i}}_k = \kappa^{\mathrm{i}}_k \exp(\mathrm{i}\varphi^{\mathrm{i}}_k)$, with $k, m \in 1\ldots N$, where $N = N_\parallel N_\perp$. For simplicity, we assume that $\Gamma^{\mathrm{o}}_k = \Gamma^{\mathrm{o}}$, $|\mu_{km}| = \mu$ and $\kappa^{\mathrm{i}}_k = \kappa^{\mathrm{i}}$, if the k-th cavity is coupled to the waveguide. While this is a good approximation, Γ^{o}_k, $|\mu_{km}|$ and κ^{i}_k can vary to some degree, e.g. due to fabrication tolerances. The phase terms φ_{km} and φ^{i}_k are defined by a chip geometry and are considered independent and identically distributed (i.i.d.). The strength of intra-cavity and the cavity-waveguide coupling is weak when fulfilling

$$|\mathrm{i}\boldsymbol{\omega} \odot \boldsymbol{a}| \gg \left| -\frac{1}{2}\boldsymbol{\Gamma}^{\mathrm{e}} \odot \boldsymbol{a} + (\hat{M}^\mu + \hat{M}^\kappa)\boldsymbol{a} + \boldsymbol{K}^{\mathrm{i}}s(t) \right|. \qquad (2)$$

Direct coupling can only induce an interaction between cavities when resonances sufficiently aligned. For that reason, unless stated otherwise, we assume $\omega_k = \omega_0$. Moreover, the direction of cavity modes should be taken into account. For instance, each microring resonator (MRR) resonance is double-degenerate with counter-propagating modes. If a MRR-based reservoir is implemented exactly as in Figure 1(a), both modes of each cavity can be excited and $\boldsymbol{a}$ would have $2N$ components. In photonic crystal (PhC) cavities, however, resonances are nondegenerate and therefore $\boldsymbol{a}$ has N components. For simplicity we transform the MRR reservoir such that only one mode of each MRR cavity is excited (see Figure 1(b)).

In a fully integrated system, cavities would also be coupled to readout waveguides. This, however, can be challenging to realize for cavities inside the grid, since the compact cavity arrangement leaves no space for such waveguides. One could consider using $N_\parallel \times 1$ or $2 \times N_\perp$ grids, where all cavities are accessible. An alternative could be the use of multiple photonic layers [22] or three-dimensional
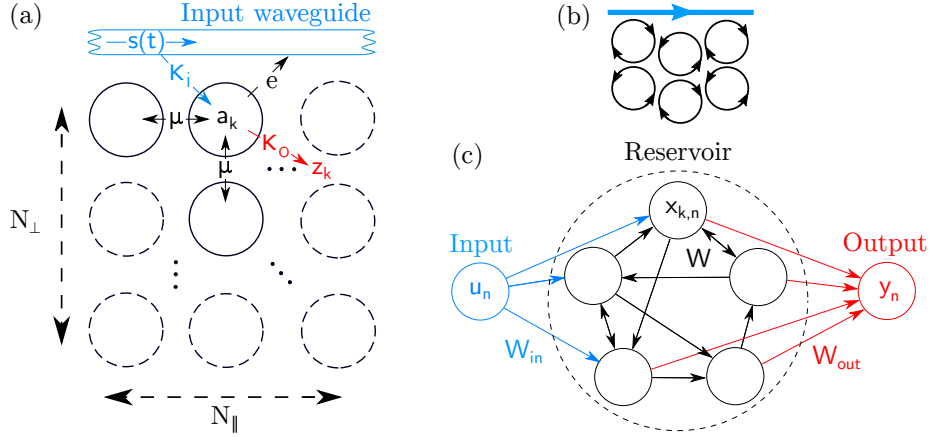
Figure 1: (a) Proposed photonic Reservoir Computer (b) MRR mode direction matching and (c) Echo State Network.

waveguides [23]. Output signals of these waveguides are $\boldsymbol{z}(t) = \boldsymbol{K}^{\mathrm{o}} \odot \boldsymbol{a}(t)$, where $\boldsymbol{K}^{\mathrm{o}}$ is a vector of output waveguide coupling coefficients, for which we assume an uniformity according to $K_k^{\mathrm{o}} = \kappa^{\mathrm{o}}$. These waveguides introduce an additional optical loss, which has to be included in $\boldsymbol{\Gamma}^{\mathrm{e}}$.

We consider three nonlinear effects commonly present in integrated photonic microresonators: TPA, FCD and the Kerr effect. TPA can be described by adding

$$-\frac{\Gamma^{\mathrm{TPA}}(a_k)}{2} a_k = -\frac{1}{2} \frac{\beta_2 c^2}{n^2 V^{\mathrm{TPA}}} |a_k|^2 a_k \tag{3}$$

to the right side of Eq. 1 [24]. Here, β_2 is a TPA coefficient, c the speed of light, n the refractive index and V^{TPA} is a nonlinear effective volume [25]

$$V^{\mathrm{TPA}} = \left(\frac{\varepsilon_0^2}{4} \int \varepsilon_{\mathrm{r}} \chi_{\mathrm{r}} \frac{2|\boldsymbol{U}|^4 + |\boldsymbol{U} \cdot \boldsymbol{U}|^2}{3} dV \right)^{-1}, \tag{4}$$

where $\varepsilon_{\mathrm{r}} = \varepsilon(\boldsymbol{r})/\varepsilon_{\max}$, $\chi_{\mathrm{r}} = \chi^{(3)}(\boldsymbol{r})/\chi_{\max}^{(3)}$ and $\boldsymbol{U}(\boldsymbol{r})$ is a cavity's normal mode, normalized such that $0.5 \int \varepsilon_0 \varepsilon_{\mathrm{r}} |\boldsymbol{U}(\boldsymbol{r})|^2 d\boldsymbol{r} = 1$.

FCD causes a microcavity's resonance to shift by $\Delta\omega_k$ as a function of free electron density N_k [26]. This effect adds a term $\mathrm{i}(\mathrm{d}\omega/\mathrm{d}N)N_k a_k$ to the right side of Eq. 1 [27]. Such free electrons could be generated via a linear photon absorption or TPA. The former is possible in direct-bandgap semiconductors, but causes high optical loss independent of the input power, unlike with TPA, where it linearly increases with the optical energy stored in a cavity. Considerable losses in systems without gain make large networks unrealistic. We therefore exclude linear photon absorption as sustainable effect producing nonlinearity and only consider TPA for the free electron generation. The density

4

of TPA-generated electrons is computed by [24]

$$\frac{\mathrm{d}N_k}{\mathrm{d}t} = -\Gamma^{\mathrm{c}}N_k + \frac{\Gamma^{\mathrm{TPA}}(a_k)}{2\hbar\omega V^{\mathrm{c}}}I_k, \tag{5}$$

where Γ^{c} is a free electron recombination rate, V^{c} is a volume across which electrons spread through a diffusion, and $I_k = |a_k|^2$ is the optical intensity inside the k-th cavity.

The Kerr effect causes a resonance frequency shift in response to a stronger electric field inside a cavity and is described by adding $\mathrm{i}(\mathrm{d}\omega/\mathrm{d}I)I_k a_k$ to the right side of Eq. 1. While the impact of the Kerr effect is typically weaker than FCD, it is still useful to be considered as an isolated case, as, unlike FCD with TPA, it provides a resonance frequency shift without an additional optical losses.

3 A Photonic Reservoir Computer based on directly coupled nonlinear cavities

A RC is often introduced as an Echo State Network (ESN) [28] (see Figure 1(c)):

$$\boldsymbol{x}_{n+1} = (1 - \alpha)\boldsymbol{x}_n + \alpha \tanh(\hat{W}\boldsymbol{x}_n + \hat{W}^{\mathrm{in}}[1; \boldsymbol{u}_{n+1}]), \tag{6}$$

where $\boldsymbol{x}_n$ and $\boldsymbol{u}_n$ are vectors of neuron activations and inputs at a time step $n \in [1, 2, \ldots T]$, α is a leaking rate, $\hat{W}$ and $\hat{W}^{\mathrm{in}}$ are a recurrent and an input weight matrices. The ESN's output is defined as $\boldsymbol{y}_n = \hat{W}^{\mathrm{out}}\boldsymbol{x}_n$, where $\hat{W}^{\mathrm{out}}$ is an output weight matrix. The readout training for software-based models is typically performed using ridge regression [28], and here we use the same approach for our simulation. Assume an ESN is being trained to solve a task with input signals $\boldsymbol{u}_n$ and the ideal target outputs $\boldsymbol{y}_n^{\mathrm{tgt}}$. First, compute update equations for T steps to obtain a set of ESN state vectors $\boldsymbol{x}_n$, which are consequently appended to create a matrix $\hat{X}$. Similarly, from $\boldsymbol{y}_n^{\mathrm{tgt}}$ one creates $\hat{Y}^{\mathrm{tgt}}$. The optimal output weight matrix is then given by

$$\hat{W}^{\mathrm{out}} = \hat{Y}^{\mathrm{tgt}}\hat{X}^{\mathrm{T}}\left(\hat{X}\hat{X}^{\mathrm{T}} + \zeta\langle|X_{ij}|^2\rangle_{ij}\hat{I}\right)^{-1}, \tag{7}$$

where $\langle\ldots\rangle_{ij}$ is an average over all matrix indices ij, ζ is a regularization constant and $\hat{I}$ is the identity matrix, superscript $(\ldots)^{\mathrm{T}}$ is the matrix transpose operation.

In the previous section, we proposed a photonic system that receives optical signal $s(t)$ as an input and provides a set of optical signals $\boldsymbol{z}(t)$. We suggest to treat this system as a reservoir computer. In the integrated photonic hardware, readout weights could be realized based on tunable attenuators with transmission w_k, phase-shifters according to φ_k and a combiner tree to implement a linear combination of $z_k(t)$ that creates the output optical signal $y(t)$:

$$y(t) = \sum_k w_k \exp(\mathrm{i}\varphi_k)z_k(t) = \boldsymbol{W}^{\mathbf{out}}\boldsymbol{z}(t). \tag{8}$$

5

This way, the conversion of optical input $s(t)$ to optical output $y(t)$ is done fully optically, as electronics are only controlling attenuators and phase shifters that are constant during a computation.

4 Linear regime

We first consider the case when the input signal is not sufficiently strong to induce a nonlinear response by the microcavities. An important value for the following discussion is the input signal's bandwidth BW that we define as the span between its spectral half-power positions.

As we will show, computational properties of such photonic reservoirs depend mostly on the microcavities' spectral characteristics. In that regard, an important consequence when directly coupling microcavities is a formation of supermodes with split resonances (see Figure 2). Consider the linear part of Eq. 1:

$$\text{diag}\left(\mathrm{i}\boldsymbol{\omega} - \frac{\boldsymbol{\Gamma}^{\mathrm{o}} + \boldsymbol{\Gamma}^{\mathrm{e}}}{2}\right) + \hat{\mathrm{M}}^{\mu} + \hat{\mathrm{M}}^{\kappa}, \tag{9}$$

eigenvalues of which are λ_k^{E}. Then the average supermode bandwidth is $\Gamma^{\mathrm{S}}/2\pi = \left\langle -2\mathrm{Re}\left(\lambda_k^{\mathrm{E}}\right)\right\rangle_k /2\pi$. Importantly, the coupling-induced supermode splitting effectively increases the reservoir's response bandwidth significantly beyond the bandwidth of a single resonator. We define the reservoir bandwidth (RBW) as the maximum distance between the network's supermodes measured at half intensity. For example, without the input waveguide, two identical independent cavities with resonances frequencies at ω_0 and coupled with rate μ become split $\omega_0 \pm |\mu|$, leading to RBW $\approx (2|\mu| + \Gamma^{\mathrm{S}})/2\pi$. For multiple indentical cavities in a chain, Eq. 9 is a tridiagonal Toeplitz matrix, eigenvalues of which are given by [29]

$$\lambda_k^{\mathrm{E}} = \left(\mathrm{i}\omega_0 + \frac{\Gamma^{\mathrm{o}}}{2}\right) + 2\sqrt{-|\mu|^2}\cos\left(\frac{k\pi}{N+1}\right), \quad k = 1\ldots N, \tag{10}$$

meaning that RBW $\approx (4|\mu| + \Gamma^{\mathrm{S}})/2\pi$. For non-trivial connectivity architectures an analytical derivation of RBW becomes non-trivial. However, in numerical simulations we found that supermode resonances stay within a $\omega_0 \pm \pi|\mu|$ interval (see Figure 2). We therefore assume that RBW $\approx (2\pi|\mu| + \Gamma^{\mathrm{S}})/2\pi$ for such directly coupled microcavity system. Furthermore, as will be shown later, in the cases we are interested in $\Gamma^{\mathrm{S}} \ll |\mu|$, and hence RBW $\approx |\mu|$.

An important computing metric of reservoirs is dimensionality D, which corresponds to the systems' degrees of freedom. Since reservoirs, and ANNs in general, perform a high-dimensional expansion of input signals [28], a higher dimensionality usually leads to a better computing capacity. To determine the dimensionality we generate a random sequence $\tilde{s}_0(t_n) = \sqrt{I_0}(t_n)\exp(\mathrm{i}\varphi_0(t_n))$, where $\sqrt{I^0}(t_n)$ and $\varphi_0(t_n)$ are randomly sampled according to a white noise distribution. Then, an 8th-order Butterworth low-pass filter with a given bandwidth BW is applied on $\tilde{s}_0(t_n)$ to obtain $s_0(t_n)$, which is then linearly interpolated to $s_0(t)$. This signal is used to modulate an optical carrier that is injected
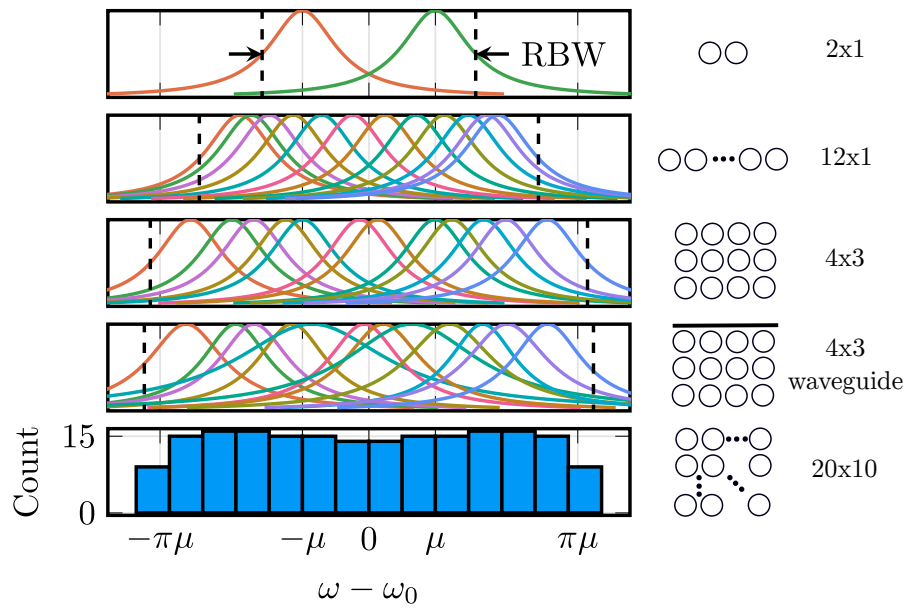
Figure 2: Supermodes of directly coupled microcavities for various grid geometries. Bottom figure shows a histogram of supermode frequencies.

7

into the reservoir as an input. Here and throughout the article the integration of differential equations is carried out using the *DifferentialEquations.jl* library [30] in the Julia programming language [31]. The photonic reservoir's dimensionality is then determined by applying the Principal Component Analysis (PCA) [32] on z_n. Importantly, for the PCA we split electric fields into a real and an imaginary part, which implies that the maximum dimensionality our reservoir can have is $2N$. It is therefore convenient to introduce a normalized dimensionality $\tilde{D} = D/2N$.

At this stage, it is beneficial to consider the supermode concept in more detail. A supermode is an independent, i.e. an orthogonal oscillation of the coupled system's electric fields. Each supermode can be considered as a virtual microcavity, that performs a pass-band filtering at its resonant frequency and bandwidth. Crucially, the coupling-induced splitting of supermode resonances makes their frequencies vary, promising high dimensionality.

To start, a supermode needs to receive an input. For that, we consider the impact of the injection BW on the reservoir's dimensionality. As we see in Figure 3(a), when BW $\ll$ RBW the dimensionality is low. In this case, a large number of supermodes are off-resonance relative to the injected field, and hence they cannot be excited. With an increasing BW, more and more supermodes are able to interact with the injected information. As a consequence, the reservoir's dimensionality increases until BW = RBW, when all supermodes become excited. Accordingly, the physics of supermodes in coupled arrays result in a necessary bandwidth matching condition for maximizing the system's dimensionality: BW $\geq$ RBW. Importantly, here the network was driven with white noise, i.e. by a signal with a flat spectrum. However, in realistic and application-relevant situations such a signal is unlikely. Certain supermodes will be excited less than others and accordingly contribute less to the networks dimensionality.

The microcavity network's dimensionality relies upon individual supermodes being driven by different aspects of an input signal, which requires supermodes to be sufficiently separated. Therefore, to quantify this metric, we consider an average supermode spacing RBW/N, normalized by the average supermode bandwidth

$$F_{\Gamma} = \frac{\text{RBW}/N}{\Gamma^{\text{S}}/2\pi}, \tag{11}$$

which can be considered as a measure of spectral span available for individual supermodes. In Figure 3(b), we see that the dimensionality is low for $F_{\Gamma} \ll 1$ and for $F_{\Gamma} > 1$ dimensionality is maximized. This threshold may change depending on system parameters, however, we find that it is typically comparable to unity. Therefore a second requirement for a high dimensionality is $F_{\Gamma} \gtrsim 1$. Again, this is not a sufficient condition, since it does not ensure that all supermodes are separated. Spacing between supermodes is typically inhomogeneous, and supermodes can potentially overlap, even if F_{Γ} is large. This effect can be mitigated through a stronger coupling to the input waveguide κ^{i}, which can shift resonances or increase the bandwidth of some supermodes, as can be seen in Figure 2 for a 4×3 reservoir with and without the input waveguide. In Figure 3(c), we see that this approach can be effective when $N_{\perp}$ is low, i.e. when

8

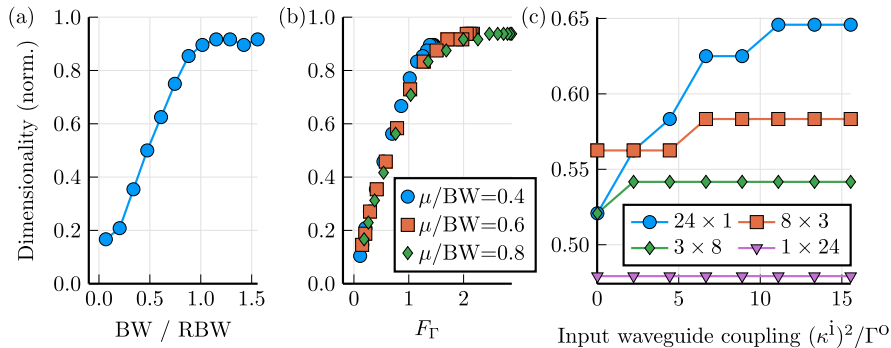overall coupling to the input waveguide is relatively stronger for the supermode splitting.

Figure 3: Dimensionality of linear photonic reservoir w.r.t. (a) input signal bandwidth, (b) average supermode spacing and (c) input waveguide coupling strength. Reservoir parameters are given in Table 1 and by default BW > RBW, but in (a) μ increased to 80 GHz to increase upper limit of dimensionality, in (b) each point corresponds to specific μ and Γ^o. (c) different grid geometries, see legend.

From F_Γ a scaling behaviour of dimensionality can be determined. Since $F_\Gamma \propto N^{-1}$, a larger number of cavities leads to a lower F_Γ, which leads to a lower $\tilde{D}$. This, however, can be compensated by lowering optical losses, as $F_\Gamma \propto (\Gamma^S)^{-1}$. Indeed, in Figure 4(a) we see that with a high μ/Γ^o (which is analogous to F_Γ), the dimensionality can reach the theoretical limit. Furthermore, dimensionality varies for a given number of cavities depending on their geometric arrangement. A minor fluctuation of dimensionality is due to a seemingly random nature of supermode formation – supermode frequencies of a 24×1 grid are different from a 12×2 grid. However, there are also outliers with a considerably lower dimensionality. In these cases $N_\perp$ is high, which allows supermodes to localize spatially further from the input waveguide. As a result, these supermodes become too weakly coupled to the input and as a consequence contribute less to the system's dimensionality (see Figure 4(b)). In addition to that, a lower $N_\perp$ allows for a higher dimensionality even at a low F_Γ, likely due to the impact of input waveguide discussed above.

5 Nonlinear regime

In order to induce a nonlinear response from the photonic system, an optical power of at least P_{NL} needs to be injected into the cavities. This power depends on material, cavity parameters and which particular nonlinearity is to be leveraged. We define P_{NL} as an input power at which a nonlinear effect becomes
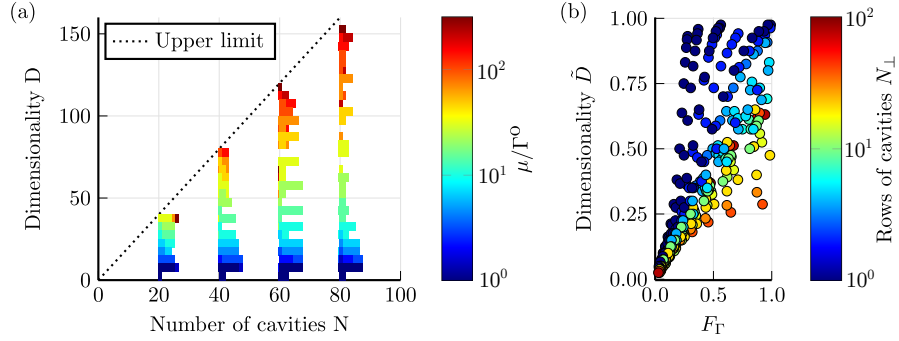
Figure 4: (a) Scalability of linear reservoir dimensionality w.r.t. optical losses. Dimensionalities of all possible combinations of $N_\parallel$ and $N_\perp$ for given N are shown as histograms. Reservoir parameters are shown in Table 1, with $\mu = 80$ GHz, reduced k_o and BW > RBW. (b) Impact of F_Γ and grid geometry on dimensionality. Both figures use the same dataset.

Table 1: Default reservoir parameters. Most parameters are omitted as they only impact the required input power.

Parameter	Value	Units	Comments
Geometry	8×3 grid		GaAs MRRs
μ	25	GHz	defines RBW
Γ^c	28.5	GHz	[33]
Q-factor	$3 \cdot 10^5$		provide a near-optimal $\Gamma^S/2\pi \approx 0.9$ GHz $\approx$ RBW$/N$
k_i^2	4π	GHz	every other cavity next to input waveguide is coupled
k_o^2	π	GHz	provides sufficiently strong $z_k(t)$

10

equal to a related linear effect. Two-photon absorption increases losses inside the cavities according to Γ^{TPA}, and it is natural to place P_{NL} where these non-linear losses start matching linear optical losses, i.e. $\Gamma^{\mathrm{TPA}} \approx \Gamma^{\mathrm{S}}$. FCD and the Kerr effect cause resonance shift $\Delta\omega$, which, in turn, is natural to be compared to an average bandwidth of supermodes, hence also Γ^{S}. Therefore, an optical loss rate is the normalization factor determining the cavity array's nonlinearity power threshold for all considered nonlinear effects.

For a single cavity injected with a monochromatic wave one can determine P_{NL} analytically (see supplementary material). The result is shown in Figure 5(a) for various integrated platforms for MRR and PhC cavities based on a data obtained from literature (see Table 2), and the horizontal axis shows an operating frequency that is BW/2. Assume a cavity with a bandwidth Γ^{S} and an electron relaxation rate Γ^{c}. This cavity can receive most optical power of an injected signal if BW $\leq \Gamma^{\mathrm{S}}/2\pi$. However, a low Γ^{S} corresponds to higher loaded Q-factor and allows for a stronger nonlinearity and, consequently, a lower P_{NL}. We therefore assume that BW $\approx \Gamma^{\mathrm{S}}/2\pi$. Additionally the scale of BW relative to Γ^{c} is also important. In the case when BW $\ll \Gamma^{\mathrm{c}}$, the generated free electrons recombine too fast and FCD becomes negligible, whereas TPA becomes the dominant nonlinearity. Then, P_{NL} is computed w.r.t. Γ^{TPA}, see dashed line in Figure 5(a). When BW is comparable to Γ^{c} FCD becomes stronger than TPA. Then, P_{NL} is computed w.r.t. $\Delta\omega$ and this case is shown as a solid line in Figure 5(a). Finally, when BW $\gg \Gamma^{\mathrm{c}}$, the electron density can't react to the changes to the input signal in time and FCD becomes irrelevant, see dotted line in Figure 5(a). Compared to MRRs, PhC cavities require less power and allow a faster operation on the same material, as a smaller mode volume corresponds to a stronger light confinement and a stronger nonlinearity, while a higher surface area enhances Γ_{c} [34].

For a network comprised of a large number of cavities, analytical treatment is not immediately tractable. We therefore resort to a numerical approach. Consider a reservoir with TPA as a nonlinearity. Since its supermodes are split due to direct coupling and, as we found in the previous section, $\Gamma^{\mathrm{S}} \ll$ RBW, a low-bandwidth excitation of this nonlinearity is not suitable. Therefore, the input changes in time and modulates Γ_k^{TPA} on a potentially individual resonator level. In order to then compare it to the scalar Γ^{S}, we consider its standard deviation across many input samples $\sigma_t[\Gamma_k^{\mathrm{TPA}}]$. Another natural option could be an average over samples, however, consider a case when TPA is dominant and $\Gamma_k^{\mathrm{TPA}}(a_k(t)) = \langle\Gamma_k^{\mathrm{TPA}}\rangle_t + \delta\Gamma_k^{\mathrm{TPA}}(a_k(t))$ with $\delta\Gamma_k^{\mathrm{TPA}}(a_k(t)) \ll \langle\Gamma_k^{\mathrm{TPA}}\rangle_t \approx \Gamma^{\mathrm{S}}$, then the loss terms in Eq. 1 for the k-th cavity is

$$\frac{\mathrm{d}a_k}{\mathrm{d}t} = \left(-\frac{\Gamma_k^{\mathrm{o}} - \Gamma_k^{\mathrm{e}}}{2} - \langle\Gamma_k^{\mathrm{TPA}}\rangle_t - \delta\Gamma_k^{\mathrm{TPA}}(a_k(t)) \right) a_k \approx \left(-\frac{\Gamma_k^{\mathrm{o}} - \Gamma_k^{\mathrm{e}}}{2} - \langle\Gamma_k^{\mathrm{TPA}}\rangle_t \right) a_k.$$

(12)

Here, $\langle\Gamma_k^{\mathrm{TPA}}\rangle_t$ effectively increases the linear loss and hence the k-th cavity is close to linear. However, averaging would still classify the cavity as a nonlinear, since $\langle\Gamma_k^{\mathrm{TPA}}(a_k(t))\rangle_t \approx \langle\Gamma_k^{\mathrm{TPA}}\rangle_t \approx \Gamma^{\mathrm{S}}$, whereas with the standard deviation $\sigma_t\left[\Gamma_k^{\mathrm{TPA}}(a_k(t))\right] = \sigma_t\left[\delta\Gamma_k^{\mathrm{TPA}}(a_k(t))\right] \ll \Gamma^{\mathrm{S}}$. We then compute $\langle\sigma_t[\Gamma_k^{\mathrm{TPA}}]\rangle_k$ to

obtain a scalar TPA rate for a given input power P. Repeating this process for a range of P we obtain a largely monotonic curve $P(\langle \sigma_t[\Gamma_k^{\text{TPA}}] \rangle_k)$, through which we can determine the input power necessary to induce a particular level of TPA-induced nonlinearity. In that case $P_{\text{NL}} = P(\Gamma^{\text{S}})$. For FCD and the Kerr effect the process to obtain a scalar characteristic capturing the network's relevant response is identical, except that $\Delta\omega_k$ is used instead of Γ_k^{TPA}. In Figure 5(b) we see that P_{NL} scales linearly with the number of cavities, provided $N_\perp$ is limited. Indeed, for large $N_\perp$ some supermodes can localize far from the input waveguide and couple to input much weaker than those near it. This creates a nonlinearity imbalance in the reservoir and to reach $\langle \sigma_t[\Gamma_k^{\text{TPA}}] \rangle_k = \Gamma^{\text{S}}$, more optical power is needed.
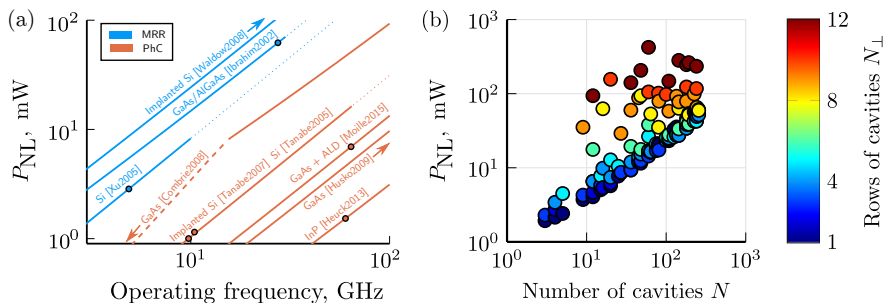
Figure 5: (a) Comparison of material platforms w.r.t. P_{NL} and an operating frequency of a single MRR or PhC cavity, details in text. Here, ALD is atomic layer deposition. For references see Table 2. (b) Scaling of P_{NL} w.r.t. the number of cavities. Reservoir parameters are given in Table 1, BW matches RBW.

Nonlinearity can result in a reservoir operating in a chaotic regime, which for most purposes is harmful to computation. Chaos corresponds to the situation where a network's state is sensitive to infinitesimal changes in the input, and hence the unavoidable presence of noise in hardware would render computations not reproducible for even perfectly identical inputs. For characterizing this effect we carry out a consistency analysis [18]. Consistency is a property describing the reproducibility of a dynamical system's (here the reservoir's) responses under an injection of slightly different input signals [19]. We generate a set of inputs in the form $s_n(t) = s_0(t) + \Delta s_n(t)$ where $s_0(t)$ and $\Delta s_n(t)$ are values sampled from a complex-valued white noise distribution, which filtered with a Butterworth 8th order pass-band filter. $s_0(t)$ is the "true input", while $\Delta s_n(t)$ emulates an impact of noise in an experiment and we attain a signal to noise ratio of 20 dB by scaling the amplitude of $\Delta s_n(t)$. Each $s_n(t)$ is injected into a separate, but identical reservoir. The absolute value of correlation between trajectories of

Table 2: References used in Figure 5.

Key	Cavity	Reference
GaAs		
Ibrahim2002	MRR	[35]
Combrie2008	L_5 PhC	[36]
Husko2009	H_0 PhC	[37]
Moille2015	$2\times$ H_0 PhC + ALD	[38]
Silicon		
Xu2005	MRR	[39]
Tanabe2005	L_3 and L_4 PhC	[34]
Implanted silicon		
Waldow2008	MRR	[40]
Tanabe2007	H_1 PhC	[41]
InP		
Ibrahim2003	MRR	[33]
Heuck2013	H_0 PhC	[42]

k-th cavity of i-th and j-th reservoirs $z_k^i(t)$ and $z_k^j(t)$ is then computed [43]

$$
\gamma_k^{ij} = \left| \frac{\mathrm{E}\left[z_k^i(z_k^j)^*\right] - \mathrm{E}\left[z_k^i\right]\mathrm{E}\left[(z_k^j)^*\right]}{\sqrt{\mathrm{E}\left[|z_k^i|^2\right]}\sqrt{\mathrm{E}\left[|z_k^j|^2\right]}} \right|,
\tag{13}
$$

where $\mathrm{E}[A]$ is expected value of A. The consistency for this pair of reservoirs is a root mean square (RMS) across all their cavities $\gamma^{ij} = \sqrt{\langle(\gamma_k^{ij})^2\rangle_k}$ and equally the consistency of system γ is the RMS of γ^{ij} of all combinations of i and j. It is typically desirable that the consistency is high [44].

In Figure 6 we show the consistency and dimensionality of reservoirs with different nonlinearities as a function of input power. Noteworthy, we find that a stronger TPA reduces the dimensionality, while the consistency remains constant and close to unity. The drop of dimensionality is a consequence of the F_Γ reduction, as increasing TPA-induced losses effectively increase Γ^S. A high consistency could be explained by the negative-feedback nature of TPA, where the system's response to a stronger input is an increased damping that often stabilizes the system. According to the data in Figure 6 FCD and the Kerr effect increase the dimensionality, but eventually the system loses consistency for too high input power. Generally, consistency decreases when $\langle\sigma_t[\Delta\omega_k]\rangle_k$ approaches Γ^S, i.e. when input power surpasses P_{NL}. The effect causing the dimensionality increase is the nonlinear shift of individual supermodes, which otherwise might overlap in the linear regime. We found that generally other system parameters have a relatively small impact compared to P_{NL}.

Interestingly, the nonlinear shift of supermode frequencies also compensates to some degree the negative impact on dimensionality caused by a mismatch between BW and RBW (see Figure 6(d)). This could be explained by the fact that a cavity might belong to multiple supermodes. When a cavity resonance shifts, it causes multiple supermodes to shift as well, even those that are not excited in a linear regime. If the shift is strong enough, these supermodes can end up inside the input bandwidth and receive input. Moreover, as BW reduces, a stronger nonlinearity is needed to move resonances into the input bandwidth, and at some point the concept of supermodes stops being applicable.
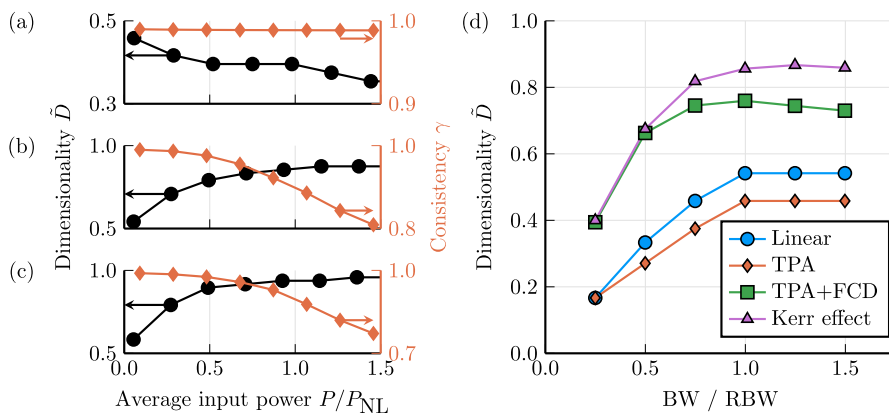
Figure 6: Effect of (a) TPA (b) TPA and FCD (c) the Kerr effect on reservoir dimensionality and consistency. Reservoir parameters are given in Table 1. (d) each point corresponds to an input power at which $\gamma = 0.98$.

To conclude, nonlinearity in general has a non-trivial effect on a reservoir. Importantly, this sensitively depends on the particular nonlinear effect that is leveraged. On one hand, dimensionality is reduced by TPA, but increased by FCD and the Kerr effect. On the other hand, using TPA the reservoir stays consistent, while FCD and the Kerr effect are harmful to consistency. However, in the case of FCD and Kerr nonlinearities, the increase of dimensionality comes before the loss of consistency. Hence, there is an interval of optimal input power that allows for a better dimensionality with a high consistency.

6   Computing performance

We now turn to evaluating the performance when applying a direct coupled cavity array reservoir to computing tests. Here, we consider the Mackey-Glass prediction task, a commonly used benchmark in the RC literature, and the optical signal recovery task that demonstrates a relevant practical application of such a system [1].

14

6.1 Mackey-Glass prediction task

The Mackey-Glass equation is a first-order time-delayed differential equation:

$$\frac{\mathrm{d}\xi}{\mathrm{d}t} = \frac{\alpha\xi(t-\tau)}{1+\xi(t-\tau)^g} - \gamma\xi(t), \tag{14}$$

where we choose $\tau = 17$, $\alpha = 0.2$, $g = 10$ and $\gamma = 0.1$, with which $\xi(t)$ exhibits a moderately chaotic behaviour [45]. Integrating Eq. 14 with the Euler method with a timestep of 0.17 and downsampling the result with a ratio of $3/0.17$, we obtain a discrete series $\xi(t_n)$. The task is to predict $y_n^{\mathrm{tgt}} = \xi(t_{n+\delta})$ with $\xi(t_n)$ as an input and a given $\delta > 0$. Both, memory and nonlinearity are required for a successful prediction, which made this task a popular RC benchmark. The prediction performance is measured with the Normalized Root Mean Square Error (NRMSE) [28]:

$$\mathrm{NRMSE}\left(\boldsymbol{y}, \boldsymbol{y}^{\mathrm{tgt}}\right) = \frac{1}{\sigma_n[y_n^{\mathrm{tgt}}]}\sqrt{\frac{1}{T}\sum_{n=1}^{T}\left|y_n - y_n^{\mathrm{tgt}}\right|^2}. \tag{15}$$

Nevertheless, the reservoir used is time-continuous, we then convert $\xi(t_n)$ into a continuous signal by a zero-hold interpolation. This signal is used to modulate an amplitude of optical carrier that is injected into reservoir as an input. To find the output $z_k(t)$ are sampled at t_n. Due to the nature of $\xi(t_n)$, the Fourier spectrum of optical input is uneven (see Figure 7). As a result, the excitation of supermodes is not uniform with a non-trivial impact on the reservoir. To visualize the impact of nonlinearities, we consider a linearized
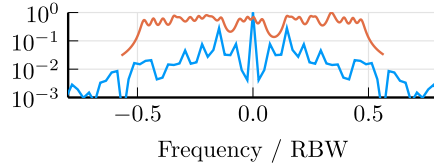

Figure 7: The normalized Fourier transform of input signal (blue) with reservoir resonances overlaid (orange). Here BW < RBW, which goes against the conclusions of Section 4. High peaks in the input spectra correspond to a dominating harmonic ($\xi(t)$ largely oscillates, see in Figure 8(c)), however, for a close prediction, it's important to also capture high-frequency irregularities, despite their weak amplitude.

Jacobian of the reservoir

$$J_{km} = \mathrm{diag}\left(\mathrm{i}\boldsymbol{\omega} - \frac{\boldsymbol{\Gamma}^{\mathrm{o}} + \boldsymbol{\Gamma}^{\mathrm{e}}}{2}\right)_{km} + \hat{\mathrm{M}}_{km}^{\mu} + \hat{\mathrm{M}}_{km}^{\kappa} + \left[-\frac{1}{2}\frac{\beta_2\mathrm{c}^2}{\mathrm{n}^2\mathrm{V}^{\mathrm{TPA}}}\mathrm{I}_k + \mathrm{i}\frac{\mathrm{d}\omega}{\mathrm{dN}}\mathrm{N}_k + \mathrm{i}\frac{\mathrm{d}\omega}{\mathrm{dI}}\mathrm{I}_k\right]\delta_{km}, \tag{16}$$

15

where last three terms correspond to TPA, FCD and the Kerr effect, with δ_{km} as the Dirac delta. In Figure 8(d) we show trajectories of $\hat{J}$ eigenvalues of the system while driven by the input. As expected, TPA increases eigenloss and modifies eigenfrequencies, most likely due to modifications to the spatial shape of the supermodes. With FCD these modifications become notably stronger, and in the case of Kerr effect they are the dominant feature. This behaviour is notably different from ESNs [46]. The inequality of nonlinearity strength is also evident. A supermode near a Fourier peak with a strong connection to an input waveguide is subject to several times stronger nonlinearity than others. This poses a challenge for FCD and the Kerr effect as a nonlinearity concentration in a few supermodes leads the reservoir to chaos before nonlinearity in other supermodes becomes strong enough.

Besides the obvious relevance of the NRMSE for performance, the output signal power also becomes a concern in noisy systems. Even though a noise was not present in simulations, we also consider a power penalty as an additional performance metric: $L_{\mathrm{p}} = 10 \log(\langle |y(t)|^2 \rangle_t / \langle |s(t)|^2 \rangle_t)$, where $\langle \dots \rangle_t$ is an average over time.

In Figure 8(a) a considerable improvement with all three nonlinearities is seen. However, with frequency-shifting nonlinearities the performance is lost after the normalized average input power approaches unity. This is consistent with Section 5 as with a loss of consistency the training becomes ineffective. However, in Figure 8(b) we see that before the consistency is lost, the output power is higher compared to TPA, which is likely due to an increase of dimensionality (see Figure 6). For a 8×3 reservoir with TPA and FCD a NRMSE of $1.5 \cdot 10^{-1}$ was achieved (see Figure 8(c)), while for TPA $8 \cdot 10^{-2}$, considering power penalty restriction. For comparison, a performance baseline is $\mathrm{NRMSE}(u_n, u_{n+3}) \approx 1.1$, an all-optical time-multiplexed reservoir with 330 virtual neurons based on a semiconductor laser experimentally achived NRMSE of 0.23 [47], a simulated time-multiplexed reservoir based on a nonlinear MRR with 25 virtual nodes had NRMSE of $7.2 \cdot 10^{-2}$ [48]. However, an ESN with 400 neurons outperforms by a large margin, predicting approximately 20 steps ahead with a NRMSE of $1.2 \cdot 10^{-4}$ [49]. Nevertheless, the proposed reservoir performs such computation fully-optically in real time.

Simulations have shown that prediction performance improves with the number of cavities (see Figure 9) for $N_\perp = 1$ or 3 and for all three nonlinearities. With FCD and the Kerr effect $N_\perp = 1$ and $N_\perp = 3$ reservoirs have shown similar performance, while for TPA $N_\perp = 3$ was better than $N_\perp = 1$. The lowest error with an NRMSE of 0.037 and L_p of -25 dB is achieved for 48 cavities with $N_\perp = 3$. We also note the importance of the direct coupling by comparing our system with a set of independent cavities, a linear regime of which was considered by [50]. For a fair comparison, resonances of cavities were set to resonances of supermodes of a corresponding $N \times 1$ reservoir. With all nonlinearities coupled cavities provided considerably better performance, likely due to a nonlinearity-induced supermode mixing.
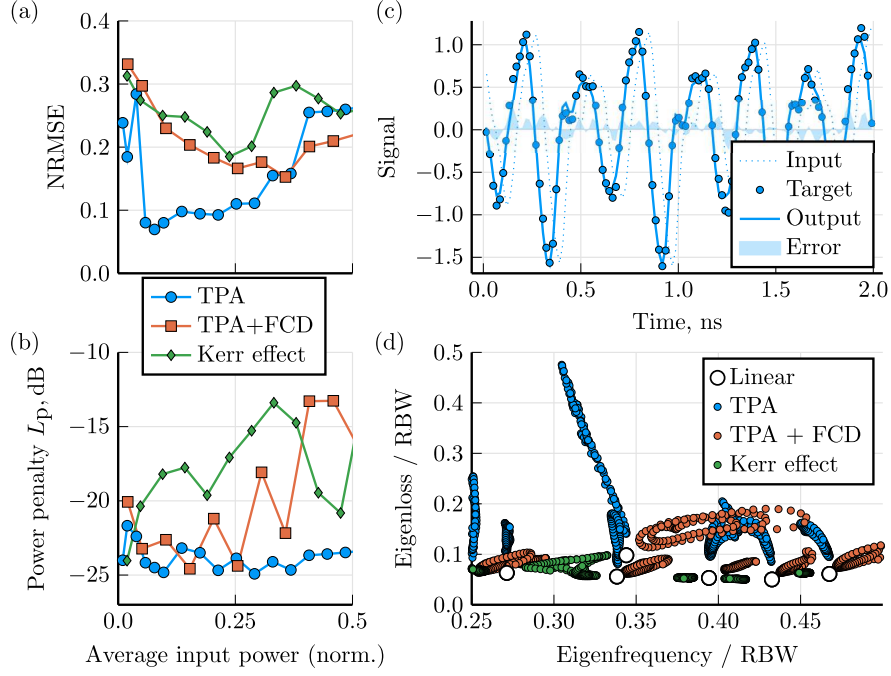
Figure 8: Impact of nonlinearities on (a) prediction error and (b) power penalty on a 3-step Mackey-Glass prediction. For all cases GaAs MRR parameters are used (see Table 1), but for TPA-only $d\omega/dN = 0$, while for Kerr effect $\beta_2 = 0$ and $d\omega/dI < 0$ is chosen arbitrarily. Regularization constant was increased until power penalty reached -25 dB. (c) 3-step prediction with TPA+FCD case at an optimal input power, here NRMSE is 0.15 and $L_{\mathrm{p}} \approx -22$ dB. (d) Trajectories of linearized Jacobian J_{km} for (c).
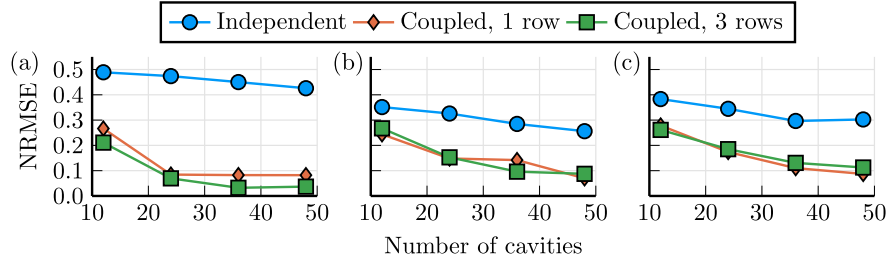
Figure 9: Scalability of 3-step prediction performance with various geometries and (a) TPA (b) TPA with FCD and (c) the Kerr effect. For each point the regularization constant was increased until power penalty reached -25 dB. Minimum NRMSE (a) 0.037 (b) 0.070 (c) 0.087.

6.2   Optical signal recovery

An optical signal propagating through an optical fiber is subject to the chromatic dispersion and the Kerr nonlinearity, among other effects. Electric field $E(z, t)$ in the fiber can be described with a nonlinear Schrödinger equation [51]:

$$\mathrm{i}\frac{\partial E}{\partial z} + \mathrm{i}\frac{\alpha_{\mathrm{loss}}}{2}E - \frac{\beta_2}{2}\frac{\partial^2 E}{\partial t^2} + \gamma|E|^2 E = 0, \tag{17}$$

where we choose $\alpha_{\mathrm{loss}} = 0.2$ dB/km, $\beta_2 = 21.7$ ps^2/km, $\gamma = 1.3$ W^{-1}/km, $E(0, t) = s_0(t)$ and $s_0(t)$ is the transmitted signal. Because of distortion, after propagating through a fiber of length L the signal at the receiver side $E(L, t) = s(t)$ differs from $s_0(t)$. Examples of the distortion are shown in Figure 10(a,b). If $s_0(t)$ is a carrier of encoded symbols, the reservoir's task is to recover them from $s(t)$.

We generate a random sequence $s_0(t_n)$ of symbols encoded with on-off keying at 25 GHz with 8 samples per bit and filter with a first order low-pass filter with $0.6 \cdot 25$ GHz cutoff frequency (see Figure 10(a)). The launch power is 10 dBm. Then Eq. 17 is solved to obtain distorted signal samples $s(t_n)$, which are linearly interpolated to $s(t)$ and injected into the reservoir. The reservoir parameters are the same as in the Mackey-Glass task (see Table 1) except for $\mu = 40$ GHz, chosen according to the $s(t)$ spectrum (see Figure 10(b)) and nonlinear effects disabled. As in the previous section, μ is increased for $N_{\parallel} \times 1$ reservoirs to equalize reservoir bandwidth. For RC training we set $y^{\mathrm{tgt}}(t) = s_0(t - \tau)$ where τ is an output delay and sample $s_0(t)$, $s(t)$ and $z_k(t)$ once per bit in its middle (see dots in Figure 10(a)). For RC testing $y(t)$ and $y^{\mathrm{tgt}}(t)$ are similarly sampled to produce y_n and y_n^{tgt} and a linear classifier C is used. Performance is measured with the Symbol Error Rate (SER): $\mathrm{SER}(y, y^{\mathrm{tgt}}) = \sum_n \mathbb{1}\left[C(y_n) \neq C(y_n^{\mathrm{tgt}})\right]/N_{\mathrm{t}}$, where N_{t} is a number of transmitted symbols. However, SER cannot be lower than $1/\min_c \left\{\sum_n \mathbb{1}[C(y_n^{\mathrm{tgt}}) = c]\right\}$, which in our case is approximately $2/N_{\mathrm{t}}$.

The presence of an output delay is motivated by group velocity dispersion that causes different input signal frequencies to arrive to a receiver at different times. By delaying $y^{\mathrm{tgt}}(t)$ we allow the reservoir to accumulate more information about the current symbol before producing an output. Simulations show that it leads to a better separation of output samples at a given power penalty. For used fiber parameters a delay of 4 bits is adequate and is used throughout the section (see Figure 11).

A larger reservoir is capable of handling a distortion from a longer fiber (see Figure 12). The performance of a $N_{\parallel} \times 1$ grid of cavities was comparable to a $N_{\parallel} \times 3$ grid and outperformed independent cavities.

7   Discussion

The proposed setup can successfully be used as a RC. However, there are notable differences that need discussion.
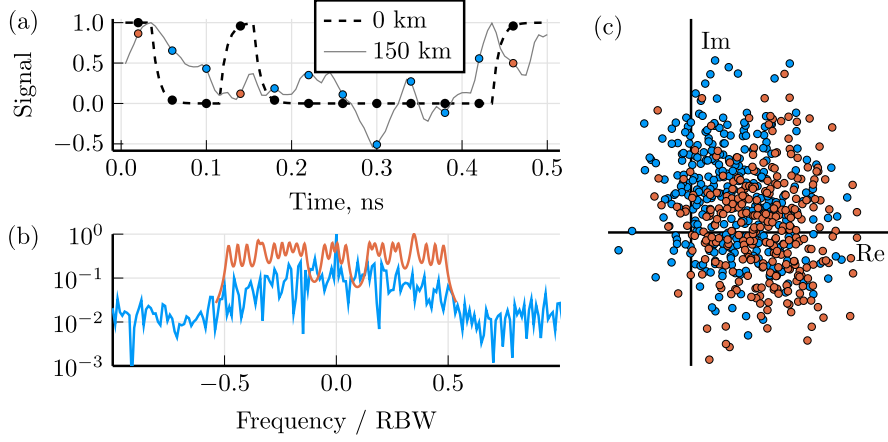
Figure 10: (a) Timetraces of real part of original signal (dashed line) and distorted signal after 150 km of optical fiber (solid line). Black dots show bit middle and act as training target, blue and orange are only for visualization and show distorted signal samples with color corresponding to each symbol. (b) Normalized Fourier spectrum of distorted signal after 150 km of fiber (blue) with 8×3 reservoir resonances overlaid (orange). (c) Constellation of distorted signal samples after 150 km of fiber, colors correspond to (a).
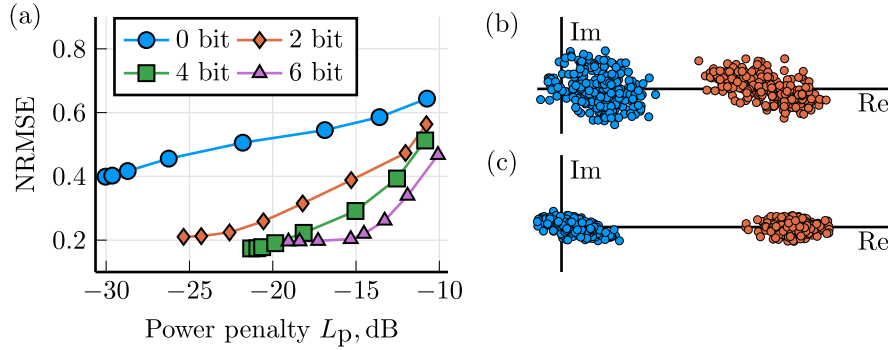
Figure 11: (a) Effect of output delay (shown on legend) on output symbol separability measured with NRMSE at a given power penalty, controlled with regularization parameter. Here $L = 150$ km and signal is processed with a 8×3 linear reservoir. (b)(c) Constellations of output symbols at $L_\mathrm{p} \approx -20$ dB with 0 and 4 bit delay respectively. Figure 10c shows constellation of input signal samples.
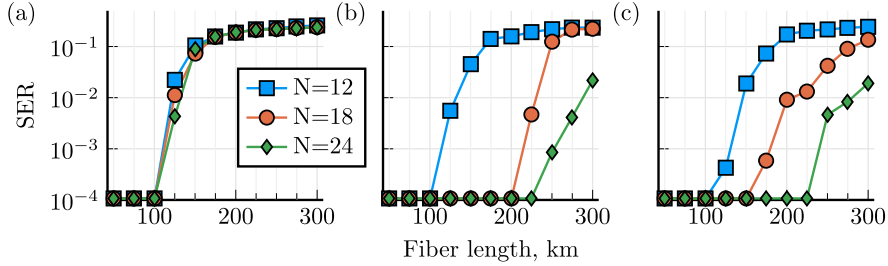
Figure 12: Performance scaling of optical signal recovery w.r.t. reservoir size. The reservoir is (a) N independent MRRs (b) $N_\parallel \times 1$ grid (c) $N_\parallel \times 3$ grid of linear MRRs. For each computation the regularization parameter is increased until the power penalty reached -25 dB.

First of all, an ESN connection matrix is typically random and sparse [28], and two neurons are unlikely to have a connection equally strong in both directions. In the photonic case if two cavities are directly coupled, the connection is reciprocal and coupling coefficients are negative complex conjugate of each other (see Figure 13).
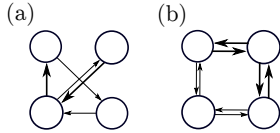

Figure 13: Neuron connectivity graph of (a) a typical ESN and (b) a photonic RC with a grid of cavities. Connection strength is represented with arrow thickness.

Moreover in ESNs a nonlinear function encompasses the input and the neuron connection terms (see Eq. 6). In our reservoir, however, a nonlinearity depends only on the current reservoir state (see Eq. 1). Nonlinearities themselves are different as well. Typically, in ESNs a sigmoid or hyperbolic tangent are used. They add a saturation of neuron activation, which is, effectively, an increase of loss. During a typical computation, ESN's linearized Jacobian eigenvalues largely move in the direction of the z-plane origin (corresponding to an increase of loss) and go back with, perhaps, slight frequency deviations [46]. In the photonic case with TPA as the only nonlinearity, the behaviour is similar, but with FCD and the Kerr effect the trajectories of eigenvalues change considerably (see Figure 8(d)). Nonetheless, all three nonlinearities positively contribute to performance (see Figure 8(a)).

In Section 4 a relation between cavity parameters and their effect on reservoir dimensionality was proposed. Whether these requirements could be met with

the available technology needs to be discussed. First, consider that $\Gamma^{\mathrm{o}} < \Gamma^{\mathrm{S}}$, $F_\Gamma \approx 1$ and RBW $\leq$ BW

$$N\frac{\Gamma^{\mathrm{o}}}{2\pi} \leq N\frac{\Gamma^{\mathrm{S}}}{2\pi} \approx NF_\Gamma\frac{\Gamma^{\mathrm{S}}}{2\pi} = \mathrm{RBW} \leq \mathrm{BW} \Rightarrow \Gamma^{\mathrm{o}} \lessapprox \frac{2\pi}{N}\mathrm{BW}. \qquad (18)$$

Another requirement demands of the reservoir to remember previous inputs: an optical timescale should be larger than an input signal timescale, i.e.

$$\frac{1}{\mathrm{BW}} < \frac{2}{\Gamma^{\mathrm{S}}} < \frac{2}{\Gamma^{\mathrm{o}}} \Rightarrow \Gamma^{\mathrm{o}} < 2\mathrm{BW}, \qquad (19)$$

which is already satisfied by Eq. 18. For example, consider a reservoir with 20 cavities. To process a signal with BW = 20 GHz, an approximate lower bound for Γ^{o} would be 2π GHz, which would correspond to an intrinsic Q-factor of $\omega_0/\Gamma^{\mathrm{o}} \approx 0.2 \cdot 10^6$. This is rather high, but not impossible. An intrinsic Q-factor of $0.7 \cdot 10^6$ was demonstrated in passive GaAs PhCs with 10^6 considered achievable [36]. In an AlGaAs MRR a Q-factor of $1.5 \cdot 10^6$ was demonstrated [52] and in a silicon PhC a Q-factor of more than $11 \cdot 10^6$ [53]. The low-loss silicon nitride allowed for even higher Q-factor of $37 \cdot 10^6$ in a MRR [54].

Next, consider a nonlinearity timescale. TPA and the Kerr effect respond almost instantaneously to a change of electric field, similar to an ESN. However, FCD is also affected by the free electron recombination rate (see Figure 5(a)), which should be comparable to a rate at which of optical signal in a cavity changes, i.e. half of its bandwidth. In our case each cavity belongs to multiple supermodes and thus its bandwidth approaches the RBW, which is close to BW. The electron recombination rate of silicon MRRs was shown to be approximately 2 GHz [55], i.e. BW $\lessapprox$ 4 GHz would be supported. The recombination can be accelerated to almost 20 GHz with a reverse-biased p-i-n junction embedded in a cavity [55]. Due to a higher surface area, in PhCs the recombination can reach 10 GHz for silicon [56] and 100 GHz for GaAs [36].

Until this point we assumed that cavities's resonances are aligned $\forall k = 1..N \; \omega_k = \omega_0$. However, this might not be the case due to fabrication tolerances. In order for directly coupled cavities to interact, resonances should be sufficiently close. A weak intercavity interaction leads to two issues, the first being a dimensionality reduction when some cavities are not coupled to the input waveguide directly (e.g. for $N_\perp > 1$). Since such cavities can only receive signals through other cavities, without an interaction they stop receiving signals and hence cannot contribute to dimensionality. This is demonstrated in Figure 14 for 8×3 and 24×1 linear MRR reservoirs. As a resonance disorder is increased, at first $\tilde{D}$ increases as the supermode overlap is alleviated, but decreases afterwards. The 24×1 reservoir is affected less than the 8×3 one, since in the former 12 cavities are coupled to the waveguide (due to MRR mode direction matching, see Figure 1(b)), but only 4 in the latter. If MRR mode direction was not considered and both MRR modes were excited, all cavities of a 24×1 MRR reservoir are coupled to a waveguide. In that case a further increase of resonance disorder simply increases F_Γ and $\tilde{D}$ will eventually reach unity. A similar outcome is expected for a PhC reservoir.
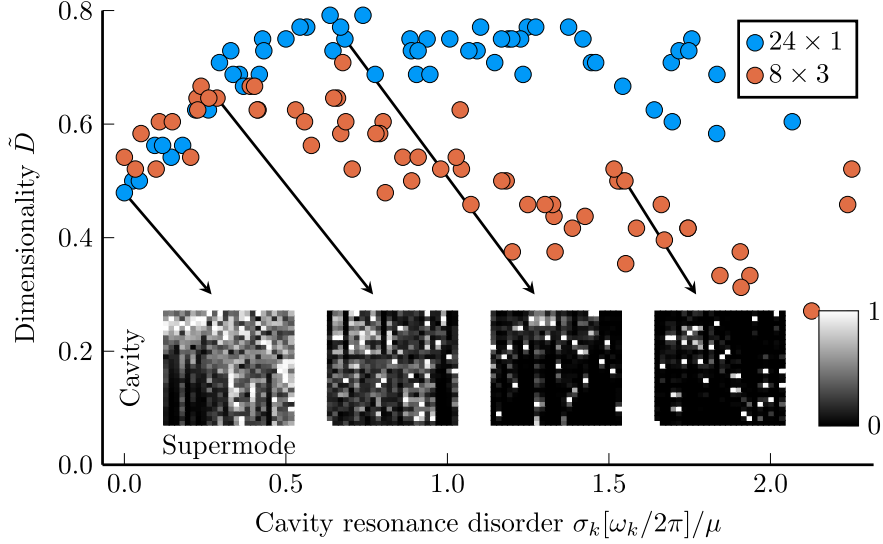
21

Figure 14: Impact of cavity resonance frequency disorder on dimensionality. Reservoir parameters given in Table 1, BW > RBW. Each dot corresponds to a $N_\parallel \times N_\perp$ (shown on legend) linear MRR reservoir with resonances $\omega_k \in \mathcal{N}(\omega_0, \sigma_\omega)$ with various σ_ω. Insets represent supermode intensity distribution over cavities $\hat{\Lambda}$.

The second issue is a weakening of supermode mixing, i.e. delocalization. When cavities interact well, supermodes span multiple cavities, otherwise, they isolate in individual cavities. Because of that, a nonlinearity of an individual cavity can only affect a single supermode. As was shown in Figure 8(d), in this case nonlinearity could not contribute to computing performance as much. Here, $N_\perp > 1$ cases could be affected to a lesser degree than $N_\perp = 1$, since for $N_\perp > 1$ each cavity has 2-4 neighbouring cavities, while only 1-2 for $N_\perp = 1$ and the probability of resonance mismatch with all neighbours is lower with more neighbours. The weakening of supermode mixing is shown in inset of Figure 14. Each heatmap represents a matrix $\hat{\Lambda}$ where Λ_{km} is an intensity of m-th supermode in the k-th cavity. For clarity, each m-th column of $\hat{\Lambda}$ is normalized to unity. We see that when $\sigma_k[\omega_k/2\pi] < 0.5\mu$, supermodes span multiple cavities, but only a few otherwise. The resonance disorder can be mitigated with the thermo-optic effect, but for more cavities it becomes challenging. In addition to that, a close proximity of cavities results in a strong thermal crosstalk. However, the resonance disorder can be less severe when cavities are spatially close [57].

8 Conclusion

We have applied a network of directly coupled nonlinear nanophotonic cavities for a reservoir computing. Such coupling allows for a lower chip footprint by several orders of magnitude, excluding readout. We have derived general conditions under which such a system attains a high dimensionality, an important property for RC. The role of nonlinear effects present in integrated microcavities, namely TPA, FCD and the Kerr effect were studied. While their impact on the reservoir dimensionality and the consistency is different, all three nonlinearities positively contributed to the performance in Mackey-Glass prediction task. The reservoir was also successful in recovery of OOK-encoded optical signal. In both cases the reservoir has shown a fully optical autonomous real-time computation in the GHz range.

Several potential improvements of this design could be proposed. For example, using multiple cavity resonances to operate on multiple channels at the same time. An another option is to overcome dimensionality limitation due to a finite Q-factor. If a reservoir has an upper limit of N cavities, one could consider using two such reservoirs with one receiving a delayed copy of input. This way the dimensionality of the setup could be expected to increase two-fold.

A Derivation of waveguide-related differential equation matrices

A mode amplitude reduces exponentially outside a resonator, therefore we can assume that each cavity-waveguide interaction happens in a restricted region between two reference planes separated by d_k (dashed lines in Figure 15) [20]. For a single cavity coupled to a waveguide the equations are known [20]:
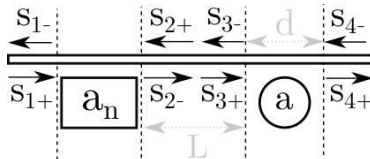

Figure 15: Scheme of waveguide-cavity interaction. A rectangle with a_n represents a row of n cavities coupled to the waveguide.

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -\left(\frac{|\kappa_1|^2}{2} + \frac{|\kappa_2|^2}{2}\right)a + \kappa_1 s_{1+} + \kappa_2 s_{2+}, \tag{20}$$

$$\begin{cases} s_{1-} = e^{-i\beta d}(s_{2+} - \kappa_2^* a) \\ s_{2-} = e^{-i\beta d}(s_{1+} - \kappa_1^* a) \end{cases}. \tag{21}$$

Omitting waveguide-induced loss and generalizing for a row of n cavities coupled to one waveguide

$$\frac{\mathrm{d}\vec{a_n}}{\mathrm{d}t} = \hat{M}^n \vec{a_n} + \vec{R_1^n} s_{1+} + \vec{R_2^n} s_{2+}, \tag{22}$$

$$\begin{pmatrix} s_{1-} \\ s_{2-} \end{pmatrix} = \begin{pmatrix} 0 & T_1^n \\ T_2^n & 0 \end{pmatrix} \begin{pmatrix} s_{1+} \\ s_{2+} \end{pmatrix} + \begin{pmatrix} \vec{Q_1^n} \\ \vec{Q_2^n} \end{pmatrix} \vec{a_n}, \tag{23}$$

where $\hat{M}^n$, $\vec{R_1^n}$ and $\vec{R_2^n}$ are assumed to be known. Adding a new cavity to the row

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \kappa_1 s_{3+} + \kappa_2 s_{4+}, \tag{24}$$

$$\begin{cases} s_{3-} = e^{-i\beta d}(s_{4+} - \kappa_2^* a) \\ s_{4-} = e^{-i\beta d}(s_{3+} - \kappa_1^* a) \\ s_{3+} = e^{-i\beta L} s_{2-} \\ s_{2+} = e^{-i\beta L} s_{3-} \end{cases}, \tag{25}$$

The goal is to find the matrix representation of a system with $n + 1$ cavities based on a system with n cavities. We first eliminate the flux between the cavities

$$\begin{cases} s_{2+} = e^{-i\beta L} e^{-i\beta d}(s_{4+} - \kappa_2^* a) \\ s_{3+} = e^{-i\beta L} \left(T_2^n s_{1+} + \vec{Q_2^n} \vec{a_n}\right) \end{cases}. \tag{26}$$

Then, cavity dynamics can be rewritten using only new inputs s_{1+} and s_{4+}:

$$\begin{cases} \frac{\mathrm{d}\vec{a_n}}{\mathrm{d}t} = \hat{M}^n \vec{a_n} + \vec{R_1^n} s_{1+} + e^{-i\beta L} e^{-i\beta d} \vec{R_2^n} s_{4+} - e^{-i\beta L} e^{-i\beta d} \vec{R_2^n} \kappa_2^* a, \\ \frac{\mathrm{d}a}{\mathrm{d}t} = e^{-i\beta L} T_2^n \kappa_1 s_{1+} + e^{-i\beta L} \vec{Q_2^n} \kappa_1 \vec{a_n} + \kappa_2 s_{4+}, \end{cases} \tag{27}$$

Consequently, new outputs are:

$$\begin{aligned} s_{1-} &= T_1^n s_{2+} + \vec{Q_1^n} \vec{a_n} = \\ &= T_1^n e^{-i\beta L} \left[e^{-i\beta d}(s_{4+} - \kappa_2^* a)\right] + \vec{Q_1^n} \vec{a_n} = \\ &= T_1^n e^{-i\beta L} e^{-i\beta d} s_{4+} + \vec{Q_1^n} \vec{a_n} - T_1^n e^{-i\beta L} e^{-i\beta d} \kappa_2^* a, \end{aligned} \tag{28}$$

$$\begin{aligned} s_{4-} &= e^{-i\beta d}(e^{-i\beta L} s_{2-} - \kappa_1^* a) = \\ &= e^{-i\beta d} \left[e^{-i\beta L} \left(T_2^n s_{1+} + \vec{Q_2^n} \vec{a_n}\right) - \kappa_1^* a\right] = \\ &= e^{-i\beta d} e^{-i\beta L} T_2^n s_{1+} + e^{-i\beta d} e^{-i\beta L} \vec{Q_2^n} \vec{a_n} - e^{-i\beta d} \kappa_1^* a. \end{aligned} \tag{29}$$

24

Therefore, the differential equation and the output matrices can be build incrementally:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\vec{a_n}\\a\end{pmatrix}=\begin{pmatrix}\hat{M}^n & -e^{-i\beta L}e^{-i\beta d}\vec{R_2^n}\kappa_2^*\\e^{-i\beta L}\vec{Q_2^n}\kappa_1 & 0\end{pmatrix}\begin{pmatrix}\vec{a_n}\\a\end{pmatrix}+$$

$$+\begin{pmatrix}\vec{R_1^n} & e^{-i\beta L}e^{-i\beta d}\vec{R_2^n}\\e^{-i\beta L}T_2^n\kappa_1 & \kappa_2\end{pmatrix}\begin{pmatrix}s_{1+}\\s_{4+}\end{pmatrix},\qquad(30)$$

$$\begin{pmatrix}s_{1-}\\s_{4-}\end{pmatrix}=\begin{pmatrix}0 & e^{-i\beta d}e^{-i\beta L}T_1^n\\e^{-i\beta d}e^{-i\beta L}T_2^n & 0\end{pmatrix}\begin{pmatrix}s_{1+}\\s_{4+}\end{pmatrix}+$$

$$+\begin{pmatrix}\vec{Q_1^n} & -e^{-i\beta L}e^{-i\beta d}T_1^n\kappa_2^*\\e^{-i\beta d}e^{-i\beta L}\vec{Q_2^n} & -e^{-i\beta d}\kappa_1^*\end{pmatrix}\begin{pmatrix}\vec{a_n}\\a\end{pmatrix}.\qquad(31)$$

B Comparing evanescent and dissipative coupling

Consider a linear case of two directly coupled lossless cavities. Their state will be described with Eq. [21]:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}a_1\\a_2\end{pmatrix}=\begin{pmatrix}0 & \mu\\-\mu^* & 0\end{pmatrix}\begin{pmatrix}a_1\\a_2\end{pmatrix},\qquad(32)$$

where μ is a coupling coefficient and * is a complex conjugation. Then $a_{1,2}\propto\exp(\pm i|\mu|t)$, i.e. no loss was introduced regardless of μ. Comparing to a similar case of waveguide coupling of two PhCs in simplified form (see A):

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}a_1\\a_2\end{pmatrix}=\begin{pmatrix}-|\kappa|^2 & |\kappa|^2\exp(i\varphi_{12})\\|\kappa|^2\exp(i\varphi_{21}) & -|\kappa|^2\end{pmatrix}\begin{pmatrix}a_1\\a_2\end{pmatrix},\qquad(33)$$

depending on φ_{12} and φ_{21} there could be an interaction ranging from the one similar to direct coupling to its total absence, i.e. light escapes from cavities to the input waveguide. In practice, however, phases are effectively random. As a result, we choose the direct coupling as a neuron connectivity method.

C   Estimation of reference power for a single cavity

Assuming a monochromatic excitation with an input power P_{NL} and a steady-state, a power that the cavity receives should be equal to a power lost:

$$P_{\mathrm{NL}} = -\frac{d|a|^2}{\mathrm{d}t} = -\frac{da}{\mathrm{d}t}a^* - \mathrm{c.c.} = \left[\Gamma^{\mathrm{S}} + \frac{\beta_2 c^2}{n^2 V^{\mathrm{TPA}}}|a|^2\right]|a|^2. \tag{34}$$

For a TPA-dominated case the reference power is found when $\Gamma^{\mathrm{TPA}}(a) = \Gamma^{\mathrm{S}}$, then

$$|a|^2 = \frac{\Gamma^{\mathrm{S}} n^2 V^{\mathrm{TPA}}}{\beta_2 c^2}. \tag{35}$$

For a FCD-dominated case in a steady state (assuming that a detuning does not impact the power absorbed by the cavity):

$$\Delta\omega = \frac{\partial\omega}{\partial N}N = \frac{\partial\omega}{\partial N}\frac{\beta_2 c^2}{n^2 V^{\mathrm{TPA}}}\frac{1}{2\hbar\omega_0 V^{\mathrm{c}}\Gamma^{\mathrm{c}}}|a|^4 = \Gamma^{\mathrm{S}}, \tag{36}$$

$$|a|^2 = \left(\frac{\Gamma^{\mathrm{S}}\Gamma^{\mathrm{c}} n^2 V^{\mathrm{TPA}} 2\hbar\omega_0 V^{\mathrm{c}}\Gamma^{\mathrm{c}}}{(\partial\omega/\partial N)\beta_2 c^2}\right)^{1/2}. \tag{37}$$

References

[1] Apostolos Argyris. "Photonic neuromorphic technologies in optical communications". In: *Nanophotonics* 11.5 (2022), pp. 897–916.

[2] Ezra Ip and Joseph M Kahn. "Compensation of dispersion and nonlinear impairments using digital backpropagation". In: *Journal of Lightwave Technology* 26.20 (2008), pp. 3416–3425.

[3] Fernando P Guiomar et al. "Mitigation of intra-channel nonlinearities using a frequency-domain Volterra series equalizer". In: *Optics express* 20.2 (2012), pp. 1360–1369.

[4] Antonio Mecozzi, Cristian Antonelli, and Mark Shtaif. "Kramers–Kronig coherent receiver". In: *Optica* 3.11 (2016), pp. 1220–1227.

[5] Christian Häger and Henry D Pfister. "Nonlinear interference mitigation via deep neural networks". In: *Optical fiber communication conference.* Optical Society of America. 2018, W3A–4.

[6] Bhavin J Shastri et al. "Principles of neuromorphic photonics". In: *arXiv preprint arXiv:1801.00016* (2017).

[7] Daniel Brunner, Miguel C Soriano, and Guy Van der Sande. "Photonic reservoir computing". In: *De Gruyter* 8 (2019), p. 19.

[8] Gyorgy Csaba and Wolfgang Porod. "Coupled oscillators for computing: A review and perspective". In: *Applied physics reviews* 7.1 (2020), p. 011302.

[9] David Marpaung, Jianping Yao, and José Capmany. "Integrated microwave photonics". In: *Nature photonics* 13.2 (2019), pp. 80–90.

[10] Yichen Shen et al. "Deep learning with coherent nanophotonic circuits". In: *Nature photonics* 11.7 (2017), pp. 441–446.

[11] Johannes Feldmann et al. "Parallel convolutional processing using an integrated photonic tensor core". In: *Nature* 589.7840 (2021), pp. 52–58.

[12] David Verstraeten et al. "An experimental unification of reservoir computing methods". In: *Neural networks* 20.3 (2007), pp. 391–403.

[13] Charis Mesaritakis, Vassilis Papataxiarhis, and Dimitris Syvridis. "Micro ring resonators as building blocks for an all-optical high-speed reservoir-computing bit-pattern-recognition system". In: *JOSA B* 30.11 (2013), pp. 3048–3055.

[14] Florian Denis-Le Coarer et al. "All-optical reservoir computing on a photonic chip using silicon-based ring resonators". In: *IEEE Journal of Selected Topics in Quantum Electronics* 24.6 (2018), pp. 1–8.

[15] Kristof Vandoorne et al. "Experimental demonstration of reservoir computing on a silicon photonics chip". In: *Nature communications* 5.1 (2014), pp. 1–6.

[16] Stijn Sackesyn et al. "Experimental realization of integrated photonic reservoir computing for nonlinear fiber distortion compensation". In: *Optics Express* 29.20 (2021), pp. 30991–30997.

[17] Joni Dambre et al. "Information processing capacity of dynamical systems". In: *Scientific reports* 2.1 (2012), pp. 1–7.

[18] Anas Skalli et al. "Computational metrics and parameters of an injection-locked large area semiconductor laser for neural network computing". In: *Optical Materials Express* 12.7 (2022), pp. 2793–2804.

[19] Atsushi Uchida, Ryan McAllister, and Rajarshi Roy. "Consistency of nonlinear system response to complex drive signals". In: *Physical review letters* 93.24 (2004), p. 244102.

[20] C Manolatou et al. "Coupling of modes analysis of resonant channel add-drop filters". In: *IEEE journal of quantum electronics* 35.9 (1999), pp. 1322–1331.

[21] Hermann Haus. "Waves and fields in optoelectronics." In: *PRENTICE-HALL, INC., ENGLEWOOD CLIFFS, NJ 07632, USA, 1984, 402* (1984).

[22] Wesley D Sacher et al. "Monolithically integrated multilayer silicon nitride-on-silicon waveguide platforms for 3-D photonic circuits and devices". In: *Proceedings of the IEEE* 106.12 (2018), pp. 2232–2245.

[23] Johnny Moughames et al. "Three-dimensional waveguide interconnects for scalable integration of photonic neural networks". In: *Optica* 7.6 (2020), pp. 640–646.

[24] Toshitsugu Uesugi et al. "Investigation of optical nonlinearities in an ultra-high-Q Si nanocavity in a two-dimensional photonic crystal slab". In: *Optics express* 14.1 (2006), pp. 377–386.

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

[26] Brian R Bennett, Richard A Soref, and Jesus A Del Alamo. "Carrier-induced change in refractive index of InP, GaAs and InGaAsP". In: *IEEE Journal of Quantum Electronics* 26.1 (1990), pp. 113–122.

[27] Qianfan Xu and Michal Lipson. "Carrier-induced optical bistability in silicon ring resonators". In: *Optics letters* 31.3 (2006), pp. 341–343.

[28] Mantas Lukoševičius. "A practical guide to applying echo state networks". In: *Neural networks: Tricks of the trade*. Springer, 2012, pp. 659–686.

[29] Devadatta Kulkarni, Darrell Schmidt, and Sze-Kai Tsui. "Eigenvalues of tridiagonal pseudo-Toeplitz matrices". In: *Linear Algebra and its Applications* 297 (1999), pp. 63–80.

[30] Christopher Rackauckas and Qing Nie. "Differentialequations.jl–a performant and feature-rich ecosystem for solving differential equations in julia". In: *Journal of Open Research Software* 5.1 (2017).

[31] Jeff Bezanson et al. "Julia: A fresh approach to numerical computing". In: *SIAM review* 59.1 (2017), pp. 65–98.

[32] Karl Pearson. "LIII. On lines and planes of closest fit to systems of points in space". In: *The London, Edinburgh, and Dublin philosophical magazine and journal of science* 2.11 (1901), pp. 559–572.

[33] Tarek A Ibrahim et al. "All-optical AND/NAND logic gates using semiconductor microresonators". In: *IEEE Photonics Technology Letters* 15.10 (2003), pp. 1422–1424.

[34] Takasumi Tanabe et al. "All-optical switches on a silicon chip realized using photonic crystal nanocavities". In: *Applied Physics Letters* 87.15 (2005), p. 151112.

[35] TA Ibrahim, V Van, and P-T Ho. "All-optical time-division demultiplexing and spatial pulse routing with a GaAs/AlGaAs microring resonator". In: *Optics Letters* 27.10 (2002), pp. 803–805.

[36] Sylvain Combrié et al. "GaAs photonic crystal cavity with ultrahigh Q: microwatt nonlinearity at 1.55 μm". In: *Optics letters* 33.16 (2008), pp. 1908–1910.

[37] Chad Husko et al. "Ultrafast all-optical modulation in GaAs photonic crystal cavities". In: *Applied Physics Letters* 94.2 (2009), p. 021111.

[38] Gregory Moille et al. "100ghz integrated all-optical switch enabled by ald". In: *arXiv preprint arXiv:1506.01609* (2015).

[39] Qianfan Xu et al. "Micrometre-scale silicon electro-optic modulator". In: *nature* 435.7040 (2005), pp. 325–327.

[40] Michael Waldow et al. "25ps all-optical switching in oxygen implanted silicon-on-insulator microring resonator". In: *Optics Express* 16.11 (2008), pp. 7693–7702.

[41] Takasumi Tanabe et al. "Fast all-optical switching using ion-implanted silicon photonic crystal nanocavities". In: *Applied Physics Letters* 90.3 (2007), p. 031115.

[42] Mikkel Heuck et al. "Heterodyne pump probe measurements of nonlinear dynamics in an indium phosphide photonic crystal cavity". In: *Applied Physics Letters* 103.18 (2013), p. 181120.

[43] Kun Il Park and M Park. *Fundamentals of probability and stochastic processes with applications to communications*. Springer, 2018.

[44] Thomas Lymburn et al. "Consistency in echo-state networks". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 29.2 (2019), p. 023118.

[45] Michael C Mackey and Leon Glass. "Oscillation and chaos in physiological control systems". In: *Science* 197.4300 (1977), pp. 287–289.

[46] Mustafa C Ozturk, Dongming Xu, and Jose C Principe. "Analysis and design of echo state networks". In: *Neural computation* 19.1 (2007), pp. 111–138.

[47] Julián Bueno et al. "Conditions for reservoir computing performance using semiconductor lasers with delayed optical feedback". In: *Optics express* 25.3 (2017), pp. 2401–2412.

[48] Giovanni Donati et al. "Microring resonators with external optical feedback for time delay reservoir computing". In: *Optics Express* 30.1 (2022), pp. 522–537.

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

[50] Michiel Hermans and Benjamin Schrauwen. "Memory in linear recurrent neural networks in continuous time". In: *Neural Networks* 23.3 (2010), pp. 341–355.

[51] GP Agrawal. *Nonlinear fiber optics*. Academic Press, 2007.

[52] Lin Chang et al. "Ultra-efficient frequency comb generation in AlGaAs-on-insulator microresonators". In: *Nature communications* 11.1 (2020), pp. 1–8.

[53] Takashi Asano et al. "Photonic crystal nanocavity with a Q factor exceeding eleven million". In: *Optics express* 25.3 (2017), pp. 1769–1777.

[54] Xingchen Ji et al. "Ultra-low-loss on-chip resonators with sub-milliwatt parametric oscillation threshold". In: *Optica* 4.6 (2017), pp. 619–624.

[55] Stefan F Preble et al. "Ultrafast all-optical modulation on a silicon chip". In: *Optics letters* 30.21 (2005), pp. 2891–2893.

[56] SW Leonard et al. "Ultrafast band-edge tuning of a two-dimensional silicon photonic crystal via free-carrier injection". In: *Physical Review B* 66.16 (2002), p. 161102.

[57] Delphin Dodane et al. "Fully embedded photonic crystal cavity with Q= 0.6 million fabricated within a full-process CMOS multiproject wafer". In: *Optics express* 26.16 (2018), pp. 20868–20877.