Increasing the nonlinear computational capacity of a spatial photonic reservoir computing system

Ian Bauwens[a,b], Krishan Harkhoe[a], Emmanuel Gooskens[b], Peter Bienstman[b], Guy Verschaffelt[a], and Guy Van der Sande[a]

[a]Applied Physics Research Group, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium
[b]Photonics Research Group, Department of Information Technology, Ghent University-IMEC, Technologiepark Zwijnaarde 126, 9052 Ghent, Belgium

ABSTRACT

Photonic reservoir computing is a neuromorphic computing framework which has been successfully used for solving various difficult and time-consuming problems. Due to its photonic nature, it offers many potential advantages such as a low-power consumption and fast processing speed. In this work, we aim to improve an already well-established design of a passive spatially distributed photonic reservoir computer, consisting of a network of waveguides connected via optical splitters and combiners. This spatially distributed architecture[1] has shown good performance on a 5-bit header recognition and an isolated spoken digit recognition task. However, this design only incorporates its nonlinearity at the photodiode in its read-out layer and is susceptible to losses within the network. Inspired by the delay-based approach to implement reservoir computing, we opt here for adding extra nonlinearity into the system to increase its nonlinear computational capacity. This is achieved by adding a single semiconductor laser as active component in an external optical delay line: light from the spatial reservoir is injected in a laser, and the optical output of the laser is then fed back to an input port of the spatial reservoir. Based on numerical simulations, we show that the nonlinear computational capacity is significantly increased by adding the feedback loop. This ultimately confirms that adding the active component can be useful for solving more complex tasks.

Keywords: delay-based reservoir computing, spatially distributed reservoir computing, temporally distributed reservoir computing, semiconductor laser, feedback

1. INTRODUCTION

Over the last years there has been a growing interest and demand in machine learning and neuromorphic computing, indicated by the recent success of ChatGPT. The training of such artificial neural networks is typically performed digitally by von Neumann architectures. However, most machine learning models contain many parameters which need to be individually optimized, resulting in a large training time and energy usage. Consequently, various platforms are being investigated as potential substitutes for the current digital machine implementations. Photonic implementations of artificial neural networks emerge as one such alternative. This interest is driven by their potential to boost computing velocity, enhance power efficiency, and leverage the inherent high parallelism offered by photonics.[2–6]

An example of this type of photonic neuromorphic architecture is given by photonic reservoir computing (RC). Reservoir computing systems, which are a form of recurrent neural networks (RNNs), comprise three distinct layers: an input layer, a reservoir, and an output layer (commonly known as the readout layer). The input layer is being used for data injection into the reservoir, while predictions are generated in the output layer using the input data. The reservoir contains many randomly interconnected nodes, and operates as a dynamic nonlinear system. The connectivity weights within this reservoir remain constant and unchanged. The sole (linear) weights

Further author information: (Send correspondence to I.B.)
: E-mail: ian.bauwens@vub.be

that are subject to training are present in the output layer, thus resulting in a substantial reduction in the overall training duration of RC when compared with conventional RNNs.

In this work, we study the computational capabilities of a spatially distributed photonic network, which can be realized through an integrated approach utilizing passive optical components as its physical nodes. We specifically employ the passive spatially distributed network, known as the four-port architecture (FPA), as introduced in Ref.[7] Although the reservoir within this network remains entirely passive, some degree of nonlinearity is introduced into the reservoir computing (RC) system through the input layer[8] and/or through the readout layer. These networks have exhibited promising performance across various tasks.[1,7,9–11] However, the nonlinearities inherent in such networks may be insufficient for certain challenging tasks, resulting in suboptimal performance.[12] We investigate the incorporation of a single nonlinear component inside the reservoir to add extra nonlinearity in the form of a single semiconductor laser (SL). This extra component can provide extra power and counter losses which occur within the network.

2. NUMERICAL IMPLEMENTATIONS

2.1 Four-port architecture

In this study, our focus is on the passive reservoir computing architecture initially introduced in Ref.[1] and subsequently improved in Ref.[7] , termed the four-port architecture (FPA). This architecture operates as a linear photonic network functioning as a multipath interferometer, comprising 16 nodes arranged in a 4×4 grid. These nodes distribute their input signals evenly across three output ports. Each node is interconnected with its neighboring nodes via multiple waveguides. Specifically, every node receives input from two adjacent nodes and transmits its signal to two other adjacent nodes. The two remaining ports of each node serve for data input and detection. The optical power at each computing node, denoted as $node_i$, is measured across all output ports using photodetectors, forming the basis for calculating weights w_i.

The discrete input data samples, denoted as u_k, are injected using a sample-and-hold method, yielding an input data stream, $u(t)$. This stream is characterized by being piecewise constant, with each segment stretched over a constant duration termed τ_M. A key parameter of this architecture is the length of the waveguides connecting two MMI couplers. This length determines an internal delay time, denoted as τ_C, between nodes within the architecture, along with a phase resulting from propagation. Therefore, the ratio $\frac{\tau_C}{\tau_M}$ holds significance as an important parameter for the architectures containing an FPA.

2.2 Four-port architecture with a semiconductor laser in the external delay line

We now combine the FPA with a single-mode SL in an external optical delay line. This architecture is shown in Fig. 1 and which we refer to as the FPA+SL. The standard input port is arbitrarily chosen to be port 1, with the remaining input ports of the architecture left unused. Regarding the output ports, one port is linked to an external optical delay line, rendering the output of one node inaccessible for the readout layer, which we refer to as the feedback port. The remaining 15 nodes are each connected to photodetectors, with signals from these detectors, denoted as $node_i$ in Fig. 1, utilized to construct the readout layer and compute linear weights (w_i) during both training and testing on unseen data.

The optical signal from the feedback port is fed into a single-mode SL using an external delay line. Subsequently, the emitted field from the SL is coupled back to the 2×1 MMI coupler, which, in turn, connects to input port 1 of the FPA, thus completing the feedback loop.

In this study, the external delay time (τ_D) of the delay line from the FPA to the laser is fixed to $\tau_D = 5\,\text{ps}$. Moreover, the waveguide lengths between the MZM and the 2×1 MMI coupler, as well as between the 2×1 MMI coupler and FPA, are both set to zero in our simulations to exclude any additional timing effects. This decision is made because such delays can be encompassed within the values of τ_C or τ_D.

For our simulations, we typically set τ_M to 30 picoseconds, unless specified otherwise, as this timeframe has been shown to yield the highest memory capacity[13,14] for delay-based reservoir computing when employing an SL. The dynamics of the single-mode laser are governed using the rate equations from Ref.[15]

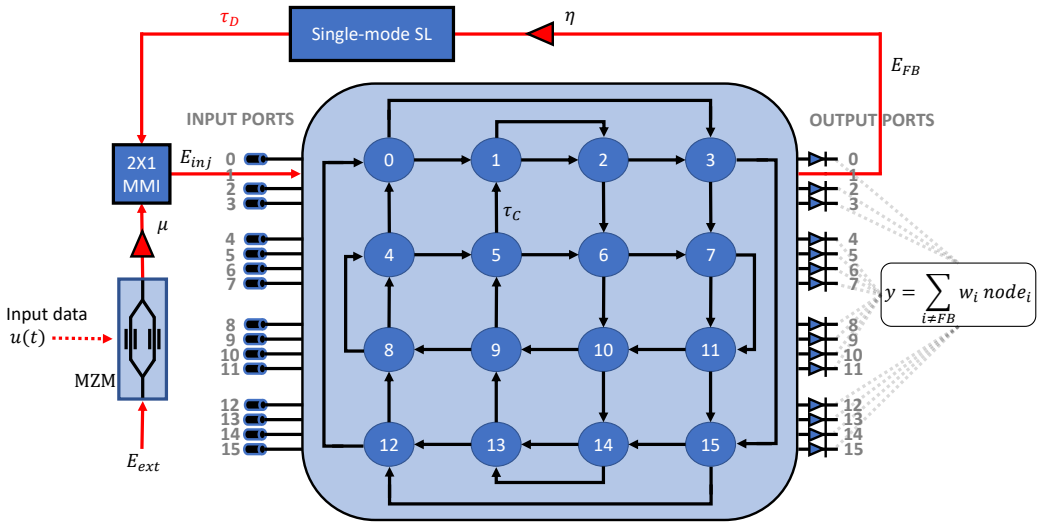

Figure 1. Illustration of the FPA+SL.

3. NUMERICAL RESULTS

We have opted to examine a task-independent metric of computational capability aimed to examine the nonlinearity of a system. Using numerical simulations, we calculate both the linear and nonlinear memory capacities of the FPA and the FPA+SL configurations.

We vary the internal delay time τ_C and the input segment duration τ_M through scanning. We employ input port 1 for data injection and as feedback port selection, we opt for output port 1. This allows for the introduction of nonlinearity early in the architecture, minimizing losses from propagation. Alternative feedback port selections may yield different outcomes in calculated memory capacity, but such considerations lie beyond the scope of this study. We calculate the memory capacity using the techniques from Ref.[16]

In Fig. 2(a)-(e), we depict the memory capacity for the initial five degrees, and the total memory capacity up to the fifth degree for the FPA in Fig. 2(f). In Fig. 2(a), we note that the linear memory capacity (i.e. degree 1, $C_{d=1}$) peaks when the internal delay time τ_C is equal to the input segment duration τ_M, i.e. $\tau_C = \tau_M$. For the second degree, $C_{d=2}$, illustrated in Fig. 2(b), we note the highest memory capacities when $\tau_C \lessapprox \tau_M$. This trend persists for the third degree, $C_{d=3}$, shown in Fig. 2(c), although the region of high memory capacity is in that case limited to $2\tau_C \lessapprox \tau_M$. For the fourth and fifth degrees, $C_{d=4}$ and $C_{d=5}$, shown in Fig. 2(d) and Fig. 2(e), we observe negligible memory capacities. The total memory capacity, shown in Fig. 2(f), shows a nearly uniform distribution.

We now compare the memory capacity of the FPA with the memory capacity of the FPA+SL. In Fig. 3, we show the memory capacity up to the fifth degree for the FPA+SL. For the linear memory capacity, $C_{d=1}$, depicted in Fig. 3(a), we observe the largest memory capacity when the internal delay time τ_C is equal to the input segment duration τ_M, i.e. $\tau_C = \tau_M$. For the second degree, $C_{d=2}$, depicted in Fig. 3(b), we note the highest memory capacities when $\tau_C \lessapprox \tau_M$. We also observe a region where the quadratic memory capacity decreases, at $\tau_C = 5\,\text{ps}$ for $\tau_M \gtrapprox 40\,\text{ps}$. The region of high cubic memory capacity, $C_{d=3}$, shown in Fig. 3(c), is found around $2\tau_C \lessapprox \tau_M$ (as was the case for the FPA). For the fourth and fifth memory capacity degrees, $C_{d=4}$ and $C_{d=5}$, shown in Fig. 3(d) and 3(e), we observe a region where these higher order memory capacities are not negligible, as opposed to the FPA. The total memory capacity, shown in Fig. 3(f), shows again a nearly uniform distribution over all τ_C and τ_M ranges, except for $\tau_C > \tau_M$.

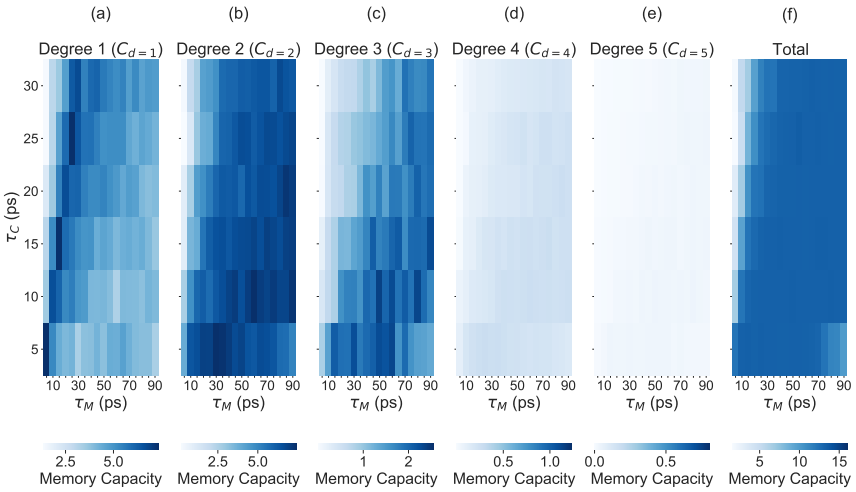

Figure 2. Memory capacity per degree for the FPA (a)-(e), and the total memory capacity (f). The internal delay time τ_C and input segment duration τ_M are varied.

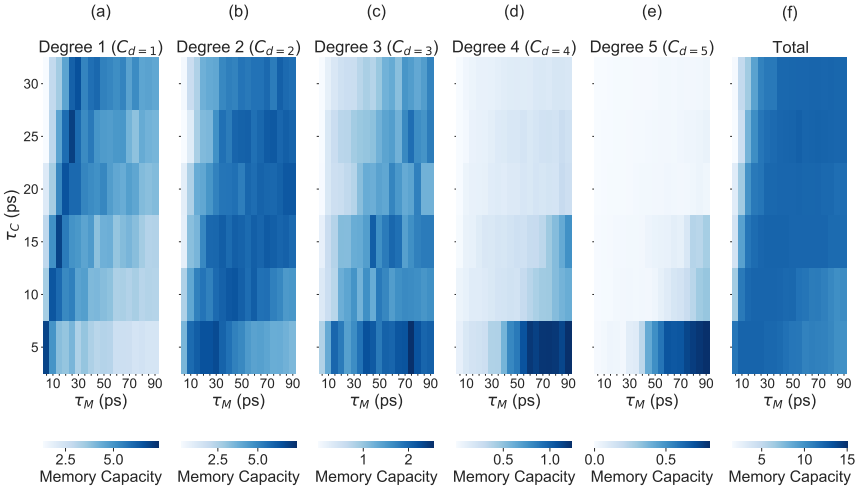

Figure 3. Memory capacity per degree for the FPA+SL (a)-(e), and the total memory capacity (f). The internal delay time τ_C and input segment duration τ_M are varied, with the external delay time fixed to $\tau_D = 5\,\text{ps}$. The injection rate is fixed to $\mu = 100\sqrt{10}\ \text{s}^{-1}$.

4. CONCLUSION

We have numerically investigated the impact of integrating a single-mode semiconductor laser to a passive spatially distributed reservoir computing system. By incorporating just a single nonlinear element into the network, we can achieve increased nonlinearity within the reservoir. This increased nonlinearity suggests that the new configuration will have better task-solving capabilities and yield improved computational performance, particularly when a more intricate nonlinearity is necessitated for handling more complex benchmark tasks.

REFERENCES

[1] Vandoorne, K., Mechet, P., Van Vaerenbergh, T., Fiers, M., Morthier, G., Verstraeten, D., Schrauwen, B., Dambre, J., and Bienstman, P., "Experimental demonstration of reservoir computing on a silicon photonics chip," *Nature communications* **5**(1), 3541 (2014).

[2] Van der Sande, G., Brunner, D., and Soriano, M. C., "Advances in photonic reservoir computing," *Nanophotonics* **6**(3), 561–576 (2017).

[3] De Lima, T. F., Shastri, B. J., Tait, A. N., Nahmias, M. A., and Prucnal, P. R., "Progress in neuromorphic photonics," *Nanophotonics* **6**(3), 577–599 (2017).

[4] Tanaka, G., Yamane, T., Héroux, J. B., Nakane, R., Kanazawa, N., Takeda, S., Numata, H., Nakano, D., and Hirose, A., "Recent advances in physical reservoir computing: A review," *Neural Networks* **115**, 100–123 (2019).

[5] Lukoševičius, M. and Jaeger, H., "Reservoir computing approaches to recurrent neural network training," *Computer science review* **3**(3), 127–149 (2009).

[6] Verstraeten, D., Schrauwen, B., d'Haene, M., and Stroobandt, D., "An experimental unification of reservoir computing methods," *Neural networks* **20**(3), 391–403 (2007).

[7] Sackesyn, S., Ma, C., Katumba, A., Dambre, J., and Bienstman, P., "A power-efficient architecture for on-chip reservoir computing," in [*Artificial Neural Networks and Machine Learning–ICANN 2019: Workshop and Special Sessions: 28th International Conference on Artificial Neural Networks, Munich, Germany, September 17–19, 2019, Proceedings 28*], 161–164, Springer (2019).

[8] Pauwels, J., Verschaffelt, G., Massar, S., and Van der Sande, G., "Distributed kerr non-linearity in a coherent all-optical fiber-ring reservoir computer," *Frontiers in Physics* **7**, 138 (2019).

[9] Vandoorne, K., Dierckx, W., Schrauwen, B., Verstraeten, D., Baets, R., Bienstman, P., and Van Campenhout, J., "Toward optical signal processing using photonic reservoir computing," *Optics express* **16**(15), 11182–11192 (2008).

[10] Vandoorne, K., Dambre, J., Verstraeten, D., Schrauwen, B., and Bienstman, P., "Parallel reservoir computing using optical amplifiers," *IEEE transactions on neural networks* **22**(9), 1469–1481 (2011).

[11] Salehi, M. R. and Dehyadegari, L., "Optical signal processing using photonic reservoir computing," *Journal of Modern Optics* **61**(17), 1442–1451 (2014).

[12] Hülser, T., Köster, F., Lüdge, K., and Jaurigue, L., "Deriving task specific performance from the information processing capacity of a reservoir computer," *Nanophotonics* **12**(5), 937–947 (2022).

[13] Harkhoe, K. and Van der Sande, G., "Task-independent computational abilities of semiconductor lasers with delayed optical feedback for reservoir computing," in [*Photonics*], **6**(4), 124, Multidisciplinary Digital Publishing Institute (2019).

[14] Bauwens, I., Harkhoe, K., Bienstman, P., Verschaffelt, G., and Van der Sande, G., "Influence of the input signal's phase modulation on the performance of optical delay-based reservoir computing using semiconductor lasers," *Optics Express* **30**(8), 13434–13446 (2022).

[15] Lenstra, D. and Yousefi, M., "Rate-equation model for multi-mode semiconductor lasers with spatial hole burning," *Optics express* **22**(7), 8143–8149 (2014).

[16] Dambre, J., Verstraeten, D., Schrauwen, B., and Massar, S., "Information processing capacity of dynamical systems," *Scientific reports* **2**(1), 1–7 (2012).