On-chip Passive Photonic Reservoir Computing with Integrated Optical Readout

Matthias Freiberger*, Andrew Katumba†, Peter Bienstman† and Joni Dambre *

* IDLab
Ghent University - imec
Technologiepark 15, 9052 Gent, Belgium
† Photonics Research Group
Ghent University - imec
Technologiepark 15, 9052 Gent, Belgium

Abstract—Photonic reservoir computing is a recent bio-inspired paradigm for signal processing. Despite first successes, the paradigm still faces challenges. We address some of these challenges and introduce our approaches to solve them. In detail, we discuss how integrated reservoirs can be scaled up by injecting multiple copies of the input. Further we introduce a new hardware-friendly training method for integrated optical readouts.

I. INTRODUCTION

Today's computational and signal processing platforms must be able to handle massive volumes of data. Reservoir computing (RC) [1], [2] initially emerged as a means to train recurrent neural networks. It embodies brain-inspired approaches to tackling the Von Neumann bottleneck, which is at the heart of the limitation of transistor-based digital implementations.

Photonic reservoir computing is an example of a hardware realization of the reservoir computing paradigm that takes advantage of the numerous optical nonlinearities for computation, as well as the inherent parallelism, large bandwidth and low power consumption of light-wave based platforms. Integrated photonic reservoirs [3] are particularly compelling, especially when implemented in the CMOS platform, as they can take advantage of low-cost mass-production technology. We have experimentally demonstrated passive reservoirs for which the nonlinearity is external to the network of components (spiral waveguides and splitters) which performs the mixing of signals [4]. For such systems, the nonlinearity required to perform complex computational tasks is naturally provided by the photodetector of the readout. Such systems are capable of processing digital optical signals at > 10 Gbps datarates, performing tasks such as header recognition, digital bitstream logical operations and signal regeneration.

While possessing numerous benefits, passive integrated photonic reservoirs are plagued by a number of issues. Among them is loss accumulation: since the interconnections between nodes are made up of spirals of a few centimeters, the material loss of for single-mode $220nm$ Si waveguides (conservatively estimated ≈ 2 dB/cm) is important. An equally significant source of signal loss occurs at combiner and splitter points. These two mechanisms constrain the useful size of the reservoirs and hence limit the complexity of tasks they can tackle.

An additional drawback of these systems is the bottleneck introduced in terms of speed and energy consumption by performing the combination of optical signals towards a desired signal in the electrical instead of the optical domain. Converting the optical states of the reservoir into the electrical domain prior to taking the weighted sum comes with a significant cost in speed and energy. This is undesired in the optical high-speed tasks which seem to be suited for these reservoirs, i.e., telecommunication applications. In order to overcome this problem, steps toward integrated optical weighting elements are currently taken in the community (for instance [5]) in order to perform signal weighting and summing entirely in the optical domain. Apart from the considerable challenges in hardware development, an integrated optical readout also calls for novel machine learning techniques. This is due to the fact that classic reservoir training algorithms assume full observability of all states which is exactly what we want to avoid for the optical signals on an integrated reservoir on chip.

In this work we discuss our approaches to addressing the issues elaborated above using numerical simulations. First, we present our findings on the study of multiple-input architectures, i.e., with the input injected into combinations of nodes. Second, we briefly introduce a new training algorithm which allows us to train passive photonic reservoirs with an integrated optical readout and using only a single high-bandwidth photodetector. In both cases, we simulate a 16-node passive photonic reservoir whose nodes are connected according to the swirl topology. Details about this topology and the simulation methodology for multiple input strategies can be found in [6].

II. RESERVOIRS WITH MULTIPLE INPUTS

Rather than pumping all the available power into a single input on one side of the reservoir as in our previous work, we have been investigating what would happen if we distribute that same available total power over a larger number of inputs, at different locations in the reservoir. At first, it would seem that this would be detrimental, because in each channel the input power is now lower, but on the other hand there is now a larger spread of the energy throughout the entire network. Our simulations have indeed suggested that this can be beneficial,

both for the power efficiency and the robustness of the system [6].

We numerically simulated a 4×4 passive photonic swirl reservoir, where we injected the input signal into combinations of 2, 4, and 8 nodes, or into all 16 nodes. Subsequently, we trained all simulated reservoirs to solve the XOR task. Figure 1 summarizes the results. Given the length of our test sequence of 10000 bits, the minimal detectable bit error rate, with a confidence level of $\approx 90\%$, is 10^{-3} [7]. The plots confirm that multiple-input reservoirs achieve significantly better power efficiency than single-input reservoirs. This is due to the increase in richness of the reservoir dynamics and the fact that signals stay longer within the reservoir. To achieve the minimal detectable error rate, a reservoir with input to all nodes requires between two and three orders of magnitude less total optical input power than the single input case. An optimized scheme in which the input is fed only to the four central nodes closely approximates this performance.
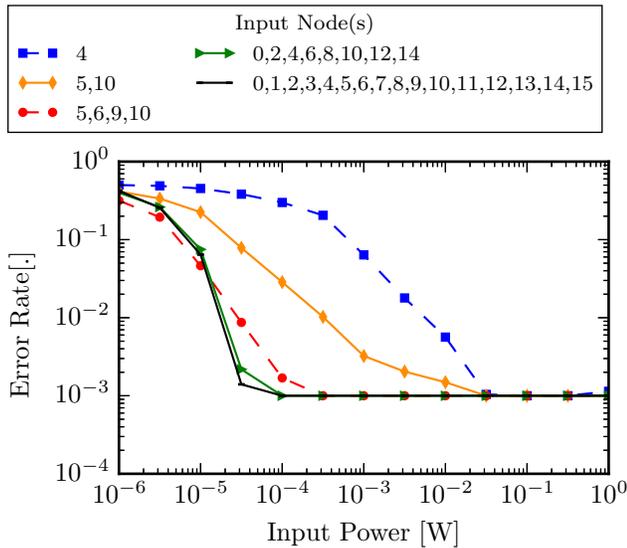

Fig. 1. Error rate vs total input power for different injection scenarios on a 4×4 passive photonic reservoir trained to solve the XOR task (node indices are ordered row by row and from left to right).

III. HARDWARE-FRIENDLY TRAINING ALGORITHMS

Regular reservoir training algorithms, e.g., ridge regression or recursive least squares, assume full observability of all states of the reservoir. For a passive photonic reservoir with on-chip readout (see Figure 2 for illustration), this full observability is only obtained if a high-speed photodetector is implemented on the chip for each reservoir state to be read out.

This multiple-detector approach is costly in terms of chip footprint and power consumption, and therefore we want to investigate an approach where the linear combination is happening in the optical domain and the result is being sent to a single photodiode. Since this loses full observability of the states, novel algorithms to train passive photonic reservoirs need to be found. A seemingly straight-forward approach is
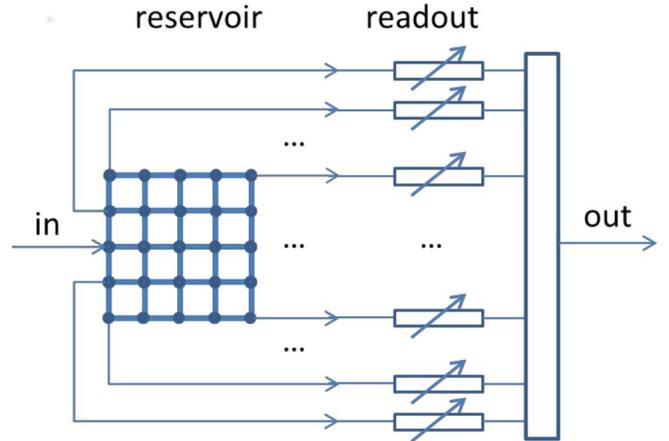

Fig. 2. Illustration of an integrated optical readout. Every node of the reservoir on chip is connected to an optical weighing element by standard silicon waveguides. After passing the optical weighting element, all weighted signals are summed by a silicon combiner structure.

to train the necessary weights for an on-chip reservoir in simulation on a digital computer, and then transfer the weights to the actual hardware. A certain increase in error when going from simulation to actual hardware is usually accepted here.

Unfortunately, in the case of passive photonic reservoirs, one deals with very high manufacturing variations of the accumulated phase in silicon waveguides and thus every single device is expected to require individual training. Therefore, simple training in simulation and using these weights for all actually fabricated devices is unlikely to yield useful results. A possible resolution to this issue is the application of a pretraining-retraining approach, where a passive photonic reservoir is pretrained in simulation and the trained weights are transferred to an actual reservoir on chip. Thereafter, the actual training error is minimized on the reservoir by fine-tuning the weights of a given integrated optical readout using a black-box optimization approach. In order to assess the feasibility of such an pretraining-retraining approach, we trained a passive 4x4 photonic swirl reservoir architecture with integrated readout to perform the 3 bit header recognition on a power-modulated digital signal fed into to the reservoir. To determine suitable values of the delay time between reservoir nodes (node interdelay) for the architecture, we trained our simulated reservoirs with increasing delay time using complex-valued ridge regression. Figure 3 shows the achieved bit error rate as a function of increasing interdelay between nodes at a input signal bitrate of 10 Gbps. Subsequently we picked the trained reservoir with a node interdelay of one bit period (100ns) and added uniform noise to the phase change induced by the waveguides of the reservoir's input connections and its connections between nodes. Figure 3b shows how the performance of the reservoir degrades in terms of bit error rate as a function of the magnitude of these emulated process variations of the reservoir (the maximal uniform phase noise

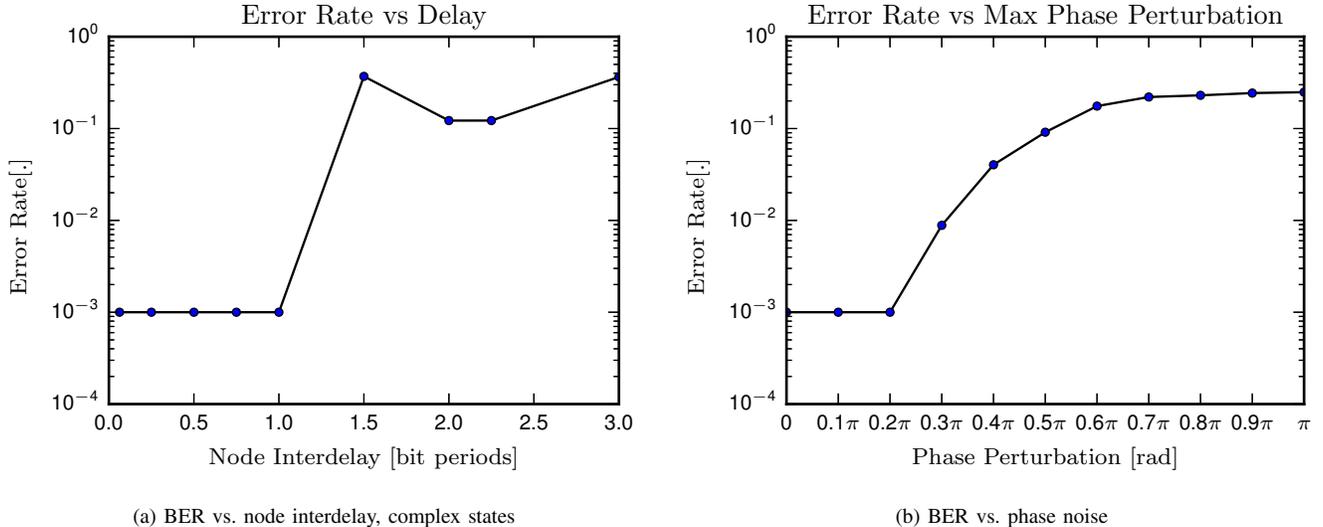(a) BER vs. node interdelay, complex states

(b) BER vs. phase noise

Fig. 3. Simulated manufacturing variations: Subfigure 3a shows the bit error rate versus the time delay between nodes (interdelay) in bit periods for a 4x4 passive swirl architecture with integrated optical readout performing 3 bit header recognition (pattern 101). The integrated optical readout is simulated by inner product multiplication of the complex reservoir states with a complex weight vector and subsequentially taking the square of the absolute value of the resulting complex output signal vector. The weight vector is trained on the complex reservoir states using complex-valued ridge regression. Input is fed to the reservoir via node 2 (row-first indexing), the minimal detectable error rate is 10^{-3}. Results are averaged over 10 reservoirs with different phase configurations for every point in the plot. Subfigure 3b shows the simulation of the impact of phase variability in silicon waveguides on the 4x4 integrated photonic reservoir with a delay time between nodes (node interdelay) of one bit period. An increasing amount of uniformly distributed phase noise was added to the waveguides feeding the input signal to the reservoir as well as to those connecting the nodes. The plot shows the calculated bit error rate as a function of maximal phase variation. All results are averaged over 10 random phase distortions per point in the plot.

applied to the waveguide connections).

As one can see, the bit error rate increases fast with the maximum amount noise added to the reservoir's waveguides, which renders a pretraining-retraining approach as proposed before challenging.

Another possibility is to exploit the weighting mechanism of the optical readout of the reservoir to read out all reservoir states through the single photodetector available at the end of the summing structure. Reading out the state variable s_i in response to the training input sequence can be simply realized by setting the weight of that state variable to 1 and all other weights to 0. By presenting the whole training input sequence to the reservoir n times, where n is the number of nodes of the reservoir, the training responses of all nodes can be collected through the single photodetector. By taking the square root of each measured power value, we can approximately invert the nonlinearity of the photodetector and obtain an estimate for the evolution of the light intensity over time at the corresponding reservoir node. However, since passive photonic reservoirs work with coherent light, it is not sufficient to know only the light intensities at the points predefined as reservoir states: we also need to know the corresponding phase of the light. While the absolute phase of the optical signal inside the reservoir is lost within the photodetection process, the relative phases between the optical state signals influence the power at the detector output. We therefore estimate the phase between two given optical signals within the integrated reservoir by obtaining the evolution of the sum of their states through time

as we apply the training signal at the reservoir's input. We now use the phase of one state signal (node) as a reference. Using the evolution of the power of sum between the reference node's signal and each other node signal, as well as the previously determined powers of all individual states, we are able to estimate the relative phase of each node signal with respect to the reference node using basic trigonometric relationships. The last stage of this calculation consists of an inverse cosine, which is is injective, in the sense that there are always two solutions within the range $[-\pi, \pi]$. To discriminate between them, we perform a third measurement between the reference node's signal and each other node's signal, now shifting the phase of the reference node's readout weight by $\frac{\pi}{2}$ and comparing with the phase estimate obtained before. As a result, the whole process requires that we feed the training sequence through the reservoir $3n - 2$ times. Under ideal conditions, this nonlinearity inversion procedure is exact. Our simulations confirm that, when exchanging the true reservoir states in our setup (Figure 3) with the states estimated through the method elaborated above, the resulting bit error plot is identical to the plot in Figure 3a. Since detector noise is neglected in our current simple model of an integrated optical readout, future work will focus on extending the proof-of-concept nonlinearity inversion approach elaborated here to a more general setting incorporating a more realistic photodetector model.

IV. CONCLUSIONS

We have presented approaches to resolve two current challenges in integrated photonic reservoir computing. On the one

hand, high losses prevent the reservoir memory to scale with increasing reservoir size. Here we have shown that feeding multiple copies of the input signal to the reservoir more energy efficient (in our numerical simulations two orders of magnitude in some cases). On the other hand, performing scaleable all-optical reservoir computing with an integrated optical readout is currently prevented by the requirement of classic reservoir training algorithms to have full state observability. We have shown proof-of-concept for a new hardware-friendly training method that handles limited observability on passive photonic reservoirs by estimating the full complex reservoir state matrix through inversion of the input nonlinearity. While our current results are demonstrated on a simple readout model, we plan to evaluate the method on more realistic models incorporating typical detector shot noise and bandwidth limitations.

REFERENCES

[1] 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. 2560, pp. 2531–2560, 2002. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6789852

[2] H. Jaeger and H. Haas, "Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication." *Science (New York, N.Y.)*, vol. 304, pp. 78–80, 2004.

[3] K. Vandoorne, J. Dambre, D. Verstraeten, B. Schrauwen, and P. Bienstman, "Parallel reservoir computing using optical amplifiers." *IEEE transactions on neural networks*, vol. 22, no. 9, pp. 1469–81, 9 2011. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5966352 http://www.ncbi.nlm.nih.gov/pubmed/21803686

[4] K. Vandoorne, P. Mechet, T. Van Vaerenbergh, M. Fiers, G. Morthier, D. Verstraeten, B. Schrauwen, J. Dambre, and P. Bienstman, "Experimental demonstration of reservoir computing on a silicon photonics chip." *Nature communications*, vol. 5, p. 3541, 1 2014. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/24662967

[5] S. Abel, T. Stferle, C. Marchiori, C. Rossel, M. D. Rossell, R. Erni, D. Caimi, M. Sousa, A. Chelnokov, B. J. Offrein, and J. Fompeyrine, "A strong electro-optically active lead-free ferroelectric integrated on silicon," *Nature Communications*, vol. 4, p. 1671, 2013. [Online]. Available: http://dx.doi.org/10.1038/ncomms2695

[6] A. Katumba, M. Freiberger, P. Bienstman, and J. Dambre, "A Multiple-Input Strategy to Efficient Integrated Photonic Reservoir Computing," *Cognitive Computation*, pp. 1–8, 4 2017. [Online]. Available: http://link.springer.com/10.1007/s12559-017-9465-5

[7] M. Jeruchim, "Techniques for estimating the bit error rate in the simulation of digital communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 2, no. 1, pp. 153–170, Jan 1984.