Improving Time Series Prediction with Ensembles of Integrated Photonic Reservoirs

Matthias Freiberger[1], **Peter Bienstman**[2], **Joni Dambre**[1]
1. IDLab, Department of Electronics and Information Systems, Ghent University - imec
2. Photonics Research Group, Department of Information Technology, Ghent University - imec

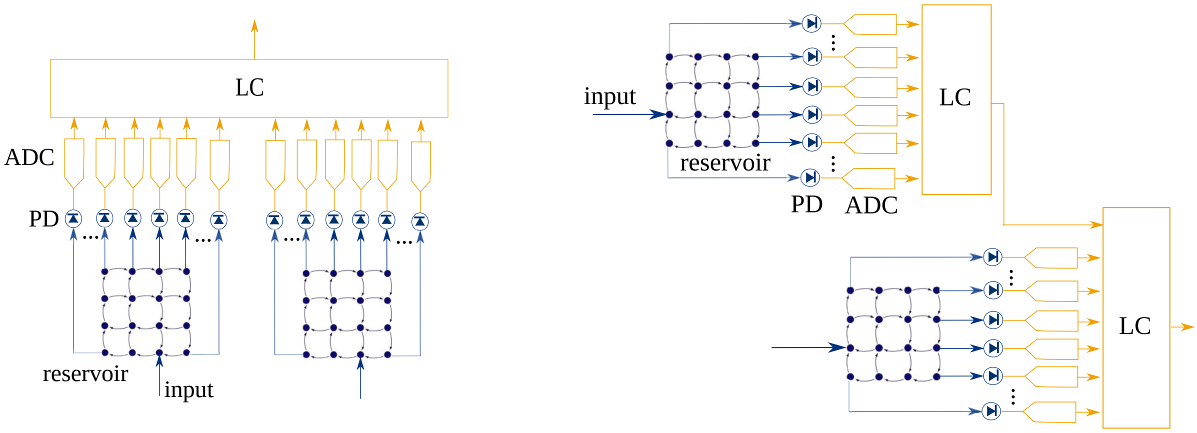
Summary. The feasible size and computational power of integrated passive photonic reservoirs is limited, amongst others, by high optical losses and wiring effort. As a step towards resolution of these issues, we propose to combine several smaller reservoirs to match or exceed the performance of a single larger one. In this abstract we investigate two possible combination techniques and evaluate their performance using the well-known Santa Fe timeseries prediction task. Our findings suggest that an ensemble of several smaller reservoirs outperforms a single reservoir with the same amount of overall nodes.

The feasible size and computational power of passive photonic reservoirs [1] is limited due to, amongst other reasons, high optical losses and wiring effort. As a step towards resolution of these issues we propose to combine several smaller reservoirs to match or exceed the performance of a single larger one. In this abstract we investigate two possible combination techniques and evaluate their performance using the well-known Santa Fe chaotic laser prediction task. We train a simple ensemble of reservoirs as well as propose a paradigm of combining reservoirs inspired by similar approaches [2, 3, 4], where we feed the output of a trained reservoir as an additional state to the classifier of a second reservoir. We refer to this technique as chaining. For details on the paradigms and the corresponding reservoir architectures, please see Figure 1.

In our simulation study, as an elementary building block of our combined designs, we use a integrated reservoir utilizing an 4x8 swirl architecture as seen in Figure 1. The reservoir uses 2x1 and 2x2 multimode interferometers as nodes, which are connected by delay lines implemented as waveguides. Delay lines are modelled to be of an approximate length of 2.35 mm with an average loss of 3dB/cm. We inject the intensity modulated input signal into 10 arbitrarily chosen input nodes. As reservoir states, we record the signal of 16 arbitrarily chosen output nodes. We combine four of these reservoirs using the ensembling and chaining approaches introduced. To ensure a fair comparison we introduce a 128 node baseline reservoir, i.e. a reservoir with four times the nodes, input and output lines of a single reservoir in our combined systems. To gain more conclusive results of the reservoirs performance we sweep the bitrate of our input signal between 20 and 40 Gbps.

As a task to evaluate the performance of our reservoirs, we use the Santa Fe timeseries prediction task [5]. We train all our readouts using ridge regression, where we perform 5-fold cross-validation to find the optimal regularization parameter for each reservoir at each bitrate. Final error measures are computed on a separate test set. We use the original train set of the Santa Fe competition for training and all remaining available data for testing, which yields train and test sets of 1000 and 9093 samples respectively. As an error measure we use the normalized mean square error (NMSE, normalized by the squared desired signal).

Figure 2 shows the NMSE obtained by all evaluated systems as a function of the input signal bitrate. As one can see, combining several reservoir poses a significant advantage over using only a single reservoir of given size for both the ensemble as well as the chaining approach. When we consider a baseline reservoir with the same overall amount of nodes as used in the combining approaches, we can observe an interesting effect: While chaining performs only slightly better than a single larger baseline reservoir, the simple ensembling approach shows significantly better performance than the baseline. While further investigation is necessary to understand this effect, a possible explanation for this effect is that several small reservoirs introduce more richness and variance in the resulting state matrix than would be possible for a single larger reservoir. To compare our results with delayed feedback approaches, we consider the work of Soriano et al. [6] who report a NMSE of 0.025 on the Santa Fe dataset for a 500 node delayed feedback reservoir computer. Our results show that this error rate can be attained by training an ensemble consisting of 4 integrated 32 node reservoirs. We conclude that combining several smaller reservoirs in place of one large reservoir poses a practical alternative to circumvent the issues arising when manufacturing large integrated photonic reservoirs. Future work will investigate, to which degree more memory and nonlinearity can be introduced into our systems by using an ensemble of smaller devices in place of a single larger one.



(a) Ensembling

(b) Chaining

Figure 1: Architectures investigated to combine reservoirs. Blue and orange colors denote optical and electrical processing respectively. PD: photodiode, ADC: analog-digital converter, LC: linear classifier. Figure 1a: Ensemble of two photonic reservoirs. Weights of all reservoirs are trained jointly. Figure 1b: Two photonic reservoirs chained together according to our proposed connection scheme.

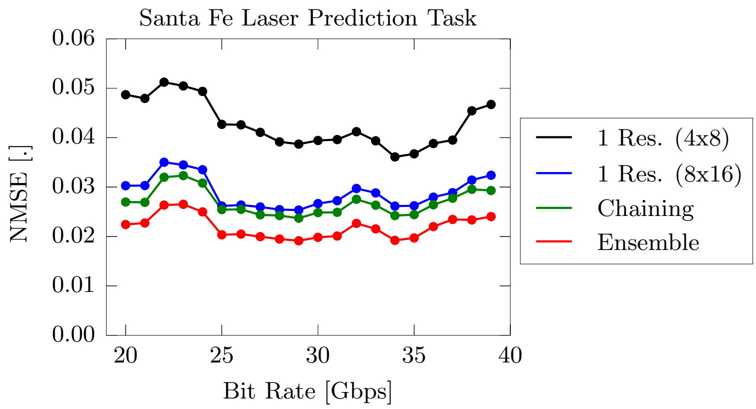


Figure 2: Normalized MSE of simulated reservoirs on the Santa Fe task as a function of bitrate for ensembling and chaining of 4 4x8 reservoirs. Comparison with 4x8 and 8x16 node baseline.

References

[1] 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*, 5(3541), 2014.

[2] L. Keuninckx. *Electronic systems as an experimental testbed to study nonlinear delay dynamics*. PhD thesis, Vrije Universiteit Brussel, 2016.

[3] C. Gallicchio and A. Micheli. Deep echo state network (deepesn): A brief survey. *arXiv preprint arXiv:1712.04323*, 2017.

[4] S. Nichele and A. Molund. Deep reservoir computing using cellular automata. *arXiv preprint arXiv:1703.02806*, 2017.

[5] A. S. Weigend and N. A. Gershenfeld. Results of the time series prediction competition at the santa fe institute. In *Neural Networks, 1993., IEEE International Conference on*, pages 1786–1793. IEEE, 1993.

[6] M. C. Soriano, S. Ortín, D. Brunner, L. Larger, C. R. Mirasso, I. Fischer, and L. Pesquera. Optoelectronic reservoir computing: tackling noise-induced performance degradation. *Optics Express*, 21(1):12–20, Jan 2013.