

Optical signal processing with a network of semiconductor optical amplifiers in the context of photonic reservoir computing

Kristof Vandoorne^a, Martin Fiers^a, David Verstraeten^b, Benjamin Schrauwen^b, Joni Dambre^b
and Peter Bienstman^a

^aPRG, Ghent University - imec, Dept. of Information Technology, Sint-Pietersnieuwstraat 41,
9000 Gent, Belgium;

^bPARIS, Ghent University, Dept. of Electronics and Information Systems,
Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium

ABSTRACT

Photonic reservoir computing is a hardware implementation of the concept of reservoir computing which comes from the field of machine learning and artificial neural networks. This concept is very useful for solving all kinds of classification and recognition problems. Examples are time series prediction, speech and image recognition. Reservoir computing often competes with the state-of-the-art. Dedicated photonic hardware would offer advantages in speed and power consumption. We show that a network of coupled semiconductor optical amplifiers can be used as a reservoir by using it on a benchmark isolated words recognition task. The results are comparable to existing software implementations and fabrication tolerances can actually improve the robustness.

Keywords: reservoir computing, integrated optics, optical neural networks, semiconductor optical amplifiers, speech recognition

1. INTRODUCTION

Reservoir Computing (RC) is a training concept for Recurrent Neural Networks (RNNs), introduced a few years ago^{1,2} coming from the field of machine learning where systems are trained based on examples. In RC a randomly initialized RNN, called the *reservoir*, is used but left untrained. The states of all the nodes of this reservoir are fed into a linear readout, which can be trained with simple and well established methods, usually linear regression. Hence, the difficulties, such as slow convergence, associated with training a full recurrent network are avoided. For several complex machine learning tasks it has been demonstrated that RC equals or outperforms other state-of-the-art techniques. An example is the prediction of the Mackey-Glass chaotic time series several of orders of magnitude better than classic methods.¹

Most reported results on reservoir computing use a (randomized) network of hyperbolic tangent or spiking neurons. However, recent work has indicated that a wide range of sufficiently high-dimensional nonlinear dynamic systems can be used as a reservoir.^{3,4} Most implementations thus far have been software based, hence the pursuit of finding a suitable hardware platform for performing the reservoir calculation. This transition offers the potential for huge power consumption savings and speed enhancement. Photonics is an interesting candidate technology for building reservoirs, because it offers a range of different nonlinear interactions working on different timescales.

In a previous paper we have shown that a network of Semiconductor Optical Amplifiers (SOAs) can be used as a reservoir on a simple signal classification task.⁵ In this paper we investigate some of the important design parameters of a photonic reservoir on a much more difficult speech recognition task. Section 2 describes the design of classical reservoirs and Sec. 3 the details of our photonic SOA reservoir. The isolated digit recognition task we use to compare classical and SOA reservoirs is discussed in Sec. 4 and the obtained results in Sec. 5. It turns out that the interconnection delay is a very important design parameter and that there is an optimal delay for this task. At this optimal delay coherent SOA reservoirs perform better than classical reservoirs.

Further author information: (Send correspondence to K.V.)

K.V.: E-mail: Kristof.Vandoorne@UGent.be, Telephone: +32 (0)9 264 3447

2. RESERVOIR COMPUTING

Echo State Networks (ESNs) are an implementation of RC where all the nodes are hyperbolic tangent functions and the network is randomly connected.¹ The reservoir state $\mathbf{x}[t]$ is updated through the following formula:

$$\mathbf{x}[t + \Delta t] = \tanh(\mathbf{W}_{in}\mathbf{u}[t] + \mathbf{W}_{res}\mathbf{x}[t]), \quad (1)$$

where $\mathbf{u}[t]$ represents the inputs, the matrix \mathbf{W}_{in} the weights for all connections from input to reservoir and \mathbf{W}_{res} is the interconnection weight matrix of the reservoir. Although the reservoir itself remains untrained (\mathbf{W}_{res} is kept fixed during training), its performance depends critically on its dynamical regime, determined by the gain and loss in the network. Optimal performance is usually obtained near the edge of stability, i.e. the region in between stable and unstable or chaotic behavior, because the system's memory is optimized there. Hence, to obtain good performance, we need to be able to tune a reservoir's dynamic regime to this edge-of-stability. A common measure for the dynamic regime is the *spectral radius*, the largest eigenvalue of the system's Jacobian, calculated at its maximal gain state (for classical hyperbolic tangent reservoirs, this corresponds to the largest eigenvalue of the network's interconnection weight matrix \mathbf{W}_{res}). The spectral radius is an indication of the stability of the network. If its value is larger than one, the network might become unstable. Tuning the spectral radius close to one often yields reservoirs with close to optimal performance.

To enhance the performance for certain tasks so called *leaky integrator neurons* are often used, where a first order recursive filter with leak rate η is added to every node:

$$\mathbf{x}[t + \Delta t] = (1 - \eta)\mathbf{x}[t] + \eta \tanh(\mathbf{W}_{in}\mathbf{u}[t] + \mathbf{W}_{res}\mathbf{x}[t]). \quad (2)$$

with $\eta \in [0, 1]$. If η equals one, we get the same situation as in equation (1). The smaller η , the higher the influence of its own past states. With this leak parameter η memory is added to all the nodes and these leaky integrator reservoirs are the classical reservoirs we compare our SOA reservoirs with.

3. PHOTONIC RESERVOIR COMPUTING

Dedicated photonic hardware offers a lot of promise, but there are some fundamental differences between a photonic and classical reservoir. For example optical information is usually encoded in changing power levels which are non-negative, in contrast to traditional RC which is real-valued amplitude based, and this makes it difficult to have negative weights and to subtract signals. In this paper we use coherent light which can be described with complex amplitudes $A = \sqrt{P} \exp(i\Phi)$ with P the power of the light and Φ the phase.

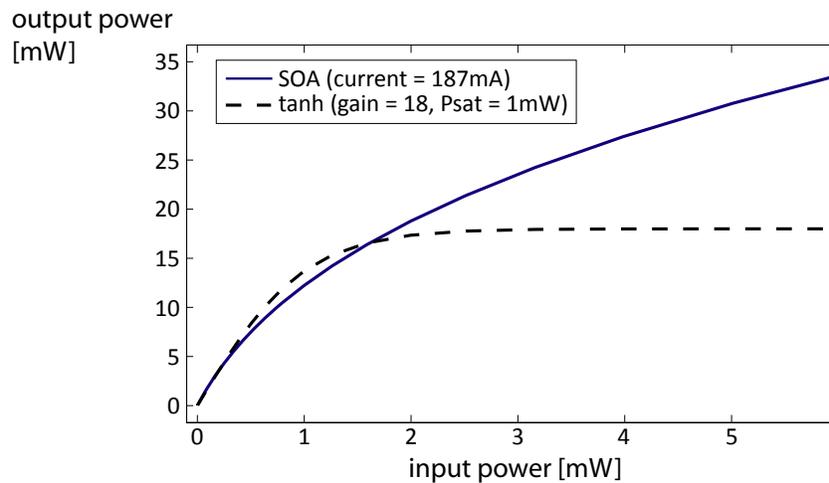


Figure 1. Comparison between a hyperbolic tangent function often used in classical reservoirs and a steady state curve of an SOA used in our photonic reservoirs.

3.1 Semiconductor Optical Amplifier

SOAs show gain saturation in their behavior and that makes them the optical device closest to the hyperbolic tangent functions used in classical reservoirs. That is the reason, besides being broadband and able to compensate losses in the network, we chose them as a first medium to verify the usefulness of photonic reservoirs. The resemblance can be seen in Fig. 1 showing the simulated steady state curve of an SOA and a scaled hyperbolic tangent. The similarity shows especially in the low-power regime (<1 mW) where the behavior is near linear. This is the regime used for the simulations in this paper. The SOA model we use is one proposed by Agrawal.⁶ It captures the most important features such as gain saturation, carrier lifetime and phase shift depending on the gain.

3.2 Topology

The topology used throughout the simulations is shown in Fig. 2 and is called a *swirl* topology as the information rotates through the network. It can be easily enlarged, while keeping the length of all connections equal. It is a nearest-neighbor topology which is easier to make, than the random networks used for classical RC, on a planar chip. We lose some freedom in connectivity that exists in software implementations, but avoid the need for a lot of crossings. Since RC was initially used with recurrent neural networks, it is important to have some feedback connections in the network, hence the swirling of the network. Having the same length for every connection makes it easier to investigate and model the influence of certain connection parameters. Every connection is defined by three parameters: the time delay Δt it introduces, the phase change $\Delta\Phi$ and the loss ΔP . The loss determines the total amount of gain and loss in the network and hence the value of the spectral radius. As was mentioned in the beginning of Sec. 3 the light is represented by complex amplitudes and since the connections influence both the phase and power of the light they act as complex weights. Therefore to incorporate the influence of coherence and interference the spectral radius has to be calculated from the now complex interconnection matrix \mathbf{W}_{res} , while also including the gain of the SOAs. Therefore the maximum gain of every SOA, i.e. at zero input (figure 1), has to be incorporated in \mathbf{W}_{res} .

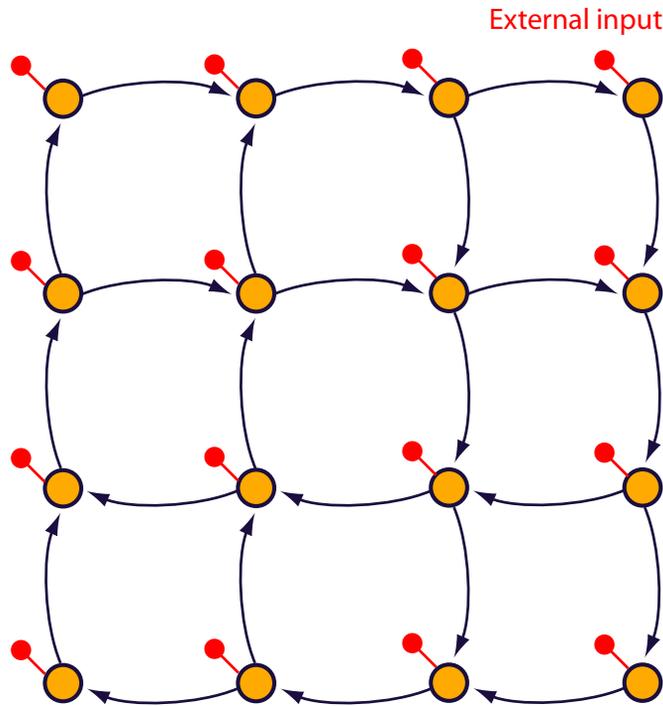


Figure 2. The swirl topology used in our simulations in a 4x4 configuration

4. SPEECH RECOGNITION

Speech recognition is very difficult to solve but reservoir computing with classical neural networks has been employed with success for this problem.⁷ The task used in this paper is the discrimination between spoken digits, the words 'zero' to 'nine', uttered by 5 female speakers. The dataset and the simulation framework for classical reservoirs is publicly available*. As is standard for speech recognition, some pre-processing of the raw speech signal is performed before it is fed into the reservoir. Often these methods involve a transformation to the frequency domain and highlighting certain frequencies typical for our ear by using some kind of ear model. The model used for the results in this paper was the Lyon ear model.⁸ We added babble noise from the NOISEX database, with a Signal-to-Noise Ratio (SNR) of 3 dB[†] to increase the complexity of the task. The performance is measured with the Word Error Rate (WER), which gives the ratio of incorrectly classified samples and the total number of samples.

Reservoir memory is related to the typical time scales of the reservoir itself. Therefore, to achieve optimal memory in a reservoir, the relevant time scales of the input signals must be adapted to those of the physical reservoir implementation. Audio signals are very slow, so we accelerated the speech signal to accord with timescales typical for the delays in our SOA network. The new durations of the digits are in the order of a few hundred ps. Hence, although we use this task to demonstrate the potential of photonic reservoir computing, we do not propose to use photonic reservoirs as a platform for standard real-time, slow audio signals.

5. RESULTS

In our experiments the input consists of 77 channels, coming from the pre-processing of the speech data. With such high-dimensional input, the number of nodes needs to be sufficiently large. All the experiments were therefore done with a rectangular 9x9 swirl network of 81 nodes. Before feeding input to the network, the different channels were mixed together in a random manner for every node and made non-negative by shifting them upwards with the minimum over all the inputs. The output of all the nodes is given to ten linear classifiers (one for every digit). These classifiers are trained on a part of the dataset and tested on a different part. The classifier with the strongest response for a certain input 'wins'.

For every simulation a sweep was done over the loss and phase change in every connection ($\Delta P, \Delta \Phi$). The loss affects the spectral radius. A result of such a sweep is shown in Fig. 3. The darker a region, the better the performance. In this case a very short delay time was used (6.25 ps) and the optimal region is only obtained for a very narrow phase region. Phase is important, since it determines the interference of the light when it combines in front of every SOA. In fact, below the instability boundary (spectral radius < 1), the reservoir is much more sensitive to phase than it is to gain.

5.1 Delay

The third parameter that defines the connections is the time delay. If we change this parameter and take each time the best result of the kind of 2D sweep described in the previous section, we get a plot as the red-circle curve in Fig. 4. First of all this figure shows that there exists an optimal delay for a given duration of the audio signal and the optimal delay corresponds to roughly half of the length of the speech signal. Second, the best classical leaky tanh reservoirs, with an optimized leak rate of 0.1, do not show this optimal delay. This suggests that delay and leak are two ways to introduce some kind of memory but combining them does not give an additional benefit. The swirl topology was also used for the classical reservoirs but they yield the same results as randomized networks for this task. The dashed line shows the best result for classical reservoirs and this means that the coherent SOA reservoir actually performs better in a broad region around the optimal delay. This makes the length of the interconnections a very important design parameter for SOA reservoirs and it will probably have to be optimized depending on the task.

*<http://snn.elis.ugent.be/rctoolbox>

†http://spib.rice.edu/spib/select_noise.html

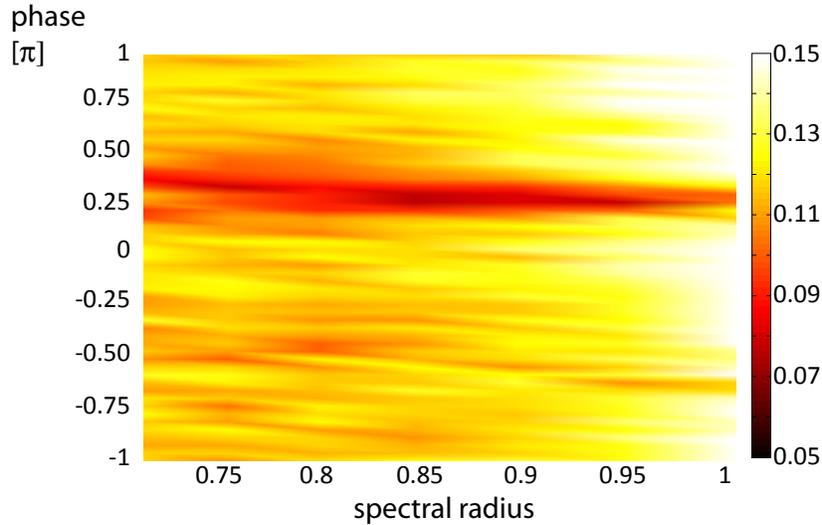


Figure 3. $(\Delta P, \Delta\Phi)$ sweep of an SOA reservoir used for isolated digit recognition.

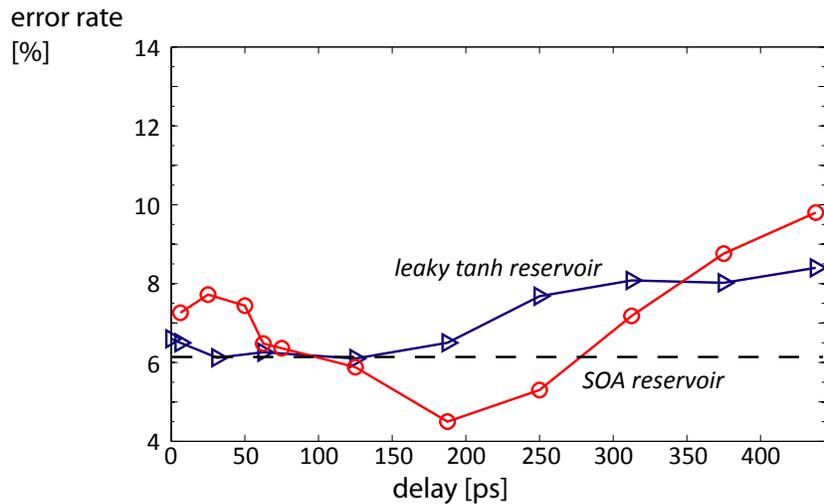


Figure 4. Result of a network of SOAs and leaky tanh reservoirs ($\eta = 0.1$) on a isolated digit speech recognition task. For both kinds of reservoir the interconnection delay is varied and the swirl topology used.

5.2 Process variations

Until now we have assumed that all connections had the same length. This condition, however, will not be fulfilled in practice since some variations on the length will inevitably exist in real hardware. We have investigated what the impact would be of small length variations on the performance. The aforementioned phase change $\Delta\Phi$ depends actually on the length d of the connections, the effective refractive index n_{eff} and the wavelength λ of the light through $\Delta\Phi = \frac{2\pi}{\lambda} n_{eff} d$. This means that variations on the length (e.g. gaussian distributed with a standard deviation of 300 nm) will result in different phase changes for different connections. Therefore the y-axis in Fig. 5 now shows the wavelength, which is one way to influence the phase change in connections. Another one would be a mechanism to alter n_{eff} . The delay used for these two simulations is 50 ps and the plots indicate that a network with variations is more robust because the optimal region is spread out over all the phases instead of being concentrated at some. The price seems a slightly worse optimal performance (7.9% versus 7.3%).

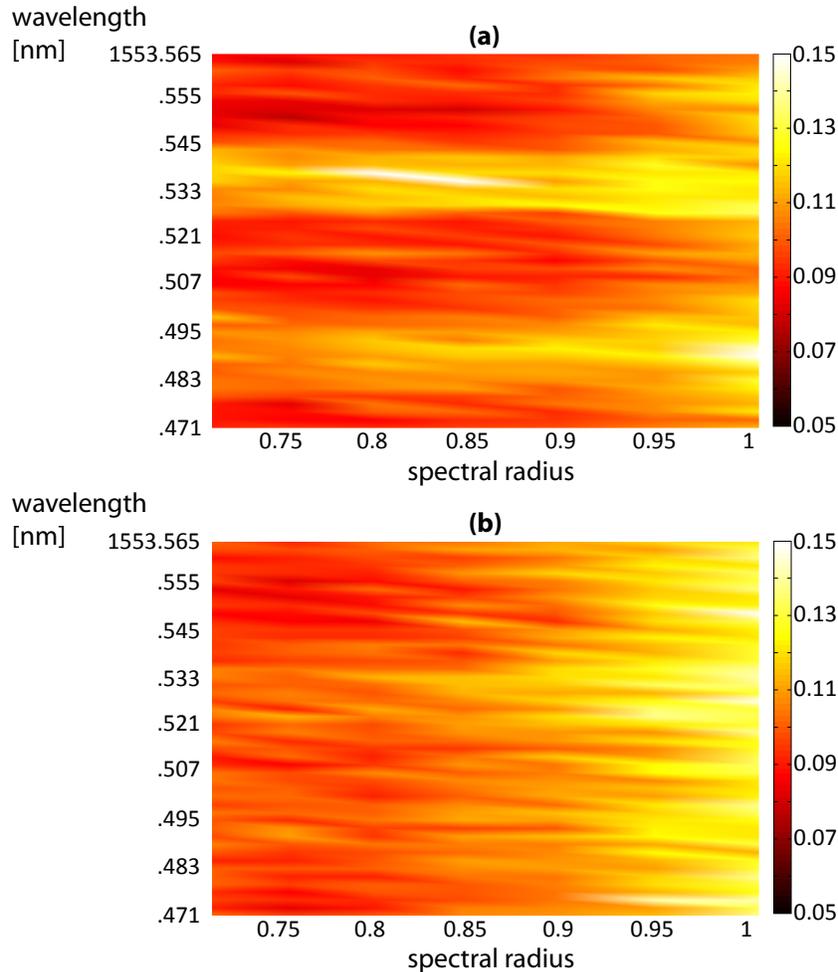


Figure 5. Result of a network of SOAs with delay 50 ps and (a) all connections having the same length; (b) length variations on all connections.

6. CONCLUSIONS

We have shown that a network of SOAs can be used as a reservoir for reservoir computing on an isolated digit recognition task and identified delay as an important design parameter. A network with optimal delay and SOAs achieves even better results on this task than classical leaky integrator tanh reservoirs. We have also shown that process variations only slightly deteriorate the optimal performance, but increase the robustness to phase changes. Future research will focus on making a small-scale hardware implementation to validate the simulation results of this paper.

ACKNOWLEDGMENTS

This work has been carried out in the framework of the IAP project Photonics@be of the Belgian Science Policy and the ERC project NaResCo. K. Vandoorne acknowledges the Special Research Fund (BOF) of Ghent University for a specialization grant.

REFERENCES

- [1] Jaeger, H. and Haas, H., “Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication,” *Science* **304**(5667), 78–80 (2004).

- [2] Maass, W., Natschläger, T., and Markram, H., “Real-time computing without stable states: A new framework for neural computation based on perturbations,” *Neural Computation* **14**(11), 2531–2560 (2002).
- [3] Verstraeten, D., Xavier-de Souza, S., Schrauwen, B., Suykens, J., Stroobandt, D., and Vandewalle, J., “Pattern classification with CNNs as reservoirs,” *Proc. International Symposium on Nonlinear Theory and its Applications (NOLTA)* (2008).
- [4] Fernando, C. and Sojakka, S., “Pattern recognition in a bucket,” *Proc. Advances in Artificial Life* **2801**, 588–597 (2003).
- [5] Vandoorne, K., Dierckx, W., Schrauwen, B., Verstraeten, D., Baets, R., Bienstman, P., and Campenhout, J. V., “Toward optical signal processing using photonic reservoir computing,” *Optics Express* **16**(15), 11182–11192 (2008).
- [6] Agrawal, G. and Olsson, N., “Self-phase modulation and spectral broadening of optical pulses in semiconductor-laser amplifiers,” *IEEE Journal of Quantum Electronics* **25**(11), 2297–2306 (1989).
- [7] Verstraeten, D., Schrauwen, B., Stroobandt, D., and Campenhout, J. V., “Isolated word recognition with the liquid state machine: a case study,” *Information Processing Letters* **95**(6), 521–528 (2005).
- [8] Lyon, R., “A computational model of filtering, detection, and compression in the cochlea,” *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* **7**, 1282–1285 (1982).