SOA-based nonlinear reservoir for echo-state networks

D.A. Ivoilov[a], A.E. Bednyakova[a], I.S. Terekhov[a], and S.K. Turitsyn[b]

[a]Novosibirsk State University, 1 Pirogova str., Novosibirsk 630090, Russia
[b]Aston Institute of Photonic Technologies, Aston University, Birmingham B4 7ET, UK

ABSTRACT

We introduce a new approach to reservoir computing (RC) in which single nonlinear device – semiconductor optical amplifier, replaces the entire nonlinear reservoir to perform computations. To study the performance of the proposed scheme, we use it for the benchmark prediction task of learning the Mackey-Glass chaotic attractor. Mildly chaotic attractor with tau = 17 and wilder chaotic behavior with tau = 30 are considered.

Keywords: semiconductor optical amplifier, recurrent neural network, reservoir computing, echo-state network, Mackey-Glass chaotic attractor

1. INTRODUCTION

Echo state network is simple and fast realization of recurrent network for dynamical systems modelling.[1–3] The idea of the echo state networks lies in two main principles, the first of which is the use of a large, random, untrainable recurrent neural network, which is called "reservoir". The second principle is to train only redout connections from reservoir to output units. Following this principles, the training of the network comes down to a simple linear regression, which is very fast and requires no iterations.

The reservoir computing (RC) post-processing of distorted signal in ultrafast optical transmission systems is a new attractive technology demonstrating high potential for efficient data decoding.[4,5] A lot of effort has gone into implementation of RC concept in nonlinear photonic systems.[6] One of the most successful implementations is delay-reservoir computer, which includes the single nonlinear node and delay line.[7] Recently, RC based on a semiconductor optical amplifier (SOA) acting as all-optical nonlinearity were proposed.[8] Here we introduce a new approach to RC in which single nonlinear device – SOA, replaces the entire nonlinear reservoir to perform computations.

2. THEORY

2.1 Semiconductor optical amplifier

The conventional model of SOA includes modifications of the power P and phase ϕ of the optical field $A = \sqrt{P}\exp(i\phi)$ by the amplification, and a differential equation for the time-dependent gain $h(t)$:[9]

$$P_{out}(t) = P_{in}(t)\exp[h(t)], \tag{1}$$

$$\phi_{out}(t) = \phi_{in}(t) - \frac{1}{2}\alpha_H h(t), \tag{2}$$

$$\frac{dh}{dt} = -\frac{h - h_0}{T_{SOA}} - \frac{P_{in}(t)}{E_{sat}}\left[\exp(h) - 1\right], \tag{3}$$

where $P_{in/out}(t)$ is the input/output power, $\phi_{in/out}(t)$ is the input/output phase of the optical signal, $\alpha_H = 5$ is the linewidth enhancement factor, $h_0 = 30$ dB is the integral small-signal gain, $T_{SOA} = 200$ ps is the gain recovery time, $E_{sat} = 8$ pJ is a characteristic saturation energy.

Transformation of temporal shape and spectrum of initially unchirped Gaussian pulse is shown in Fig. 1a,b. Pulse spectrum shifts to the red side due to SPM-induced frequency chirp imposed on the pulse as it propagates through the amplifier. Temporal dependence of the gain $h(t)$ is shown in Fig. 1c. Note, that both pulse spectrum and temporal shape are asymmetrical due to the SOA transient response. Next we exploit both SOA nonlinearity and temporal response on the order of hundreds of picoseconds to perform computations.

Further author information: (Send correspondence to D.A.I.) D.A.I.: E-mail: d.ivoilov@g.nsu.ru
A.E.B.: E-mail: anastasia.bednyakova@gmail.com

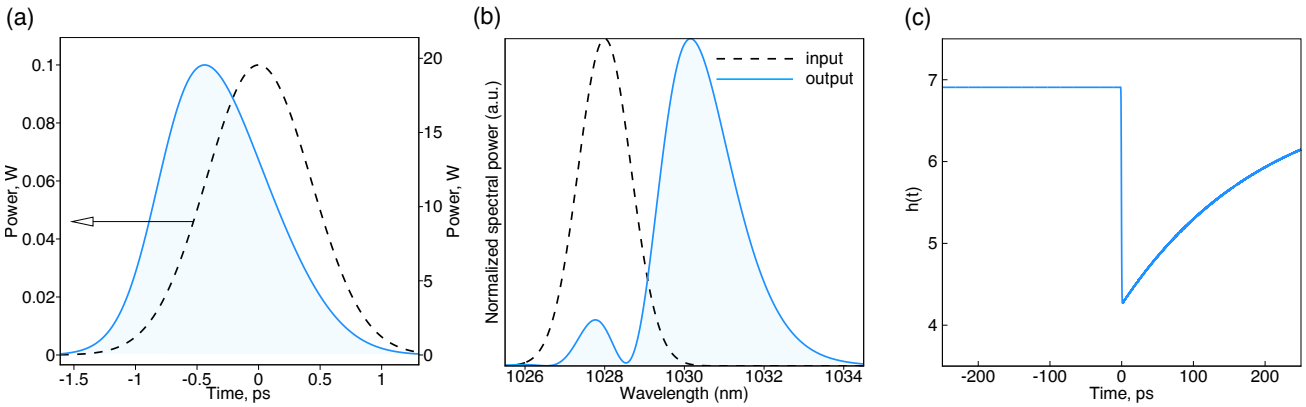

Figure 1. Pulse temporal shape (a), spectrum (b) and integral gain (c) at SOA input (dashed line) and output (blue line). $T^{in}_{FWHM} = 1$ ps, $G_0 = 30$ dB.

3. RESERVOIR COMPUTING SCHEME

The proposed scheme, shown in Fig. 2, includes input layer, SOA as a nonlinear reservoir and output layer. We use simple amplitude encoding for the input vector $x \leftrightarrow \sum_{i=1}^{N_t} \sqrt{x_i} P_0 g(t - iT)$, where T - time slot width, $g(\cdot)$ - gaussian pulse profile, P_0 - adjustable normalization coefficient that determines the amplitude level of the pulse train at SOA input. The N_c redout channels are linearly distributed in each of the N_t time slotes. The redout channels are linearly combined with the weights β_j, found during training by simple linear regression:

$$\sum_{j=1}^{Nc} W_{tj}\beta_j = \sum_{j=1}^{Nc} f_j[x^{(t)}]\beta_j = y^{(t)}, \tag{4}$$

where $\beta = (\beta^{(1)}, \beta^{(2)}, ..., \beta^{(Nc)})$ – the vector of the output weights, $x = (x^{(1)}, x^{(2),...,x^{(Nt)}})$ – the vector of training points, $y = (y^{(1)}, y^{(2)}, ..., y^{(Nt)})$ – the teacher vector, $W_{tj} = f_j[x^{(t)}]$ – the feature matrix, $f_j = P_{out}(t)$ - the redout functions, describing signal power after nonlinear transformation in SOA (see eq.(1)).

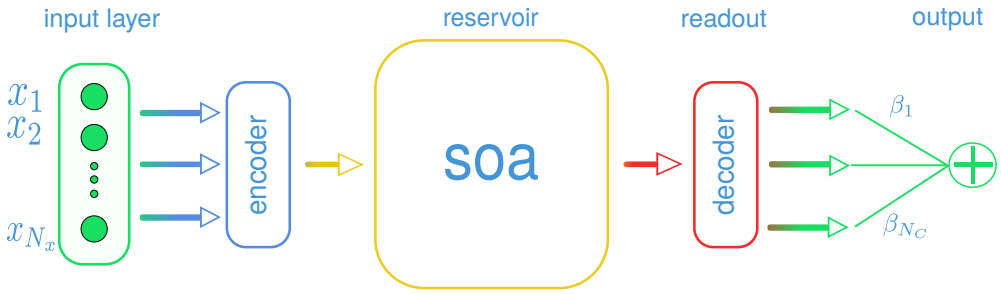

Figure 2. The basic SOA-based network architecture for reservoir computing.

4. RESULTS

To study the performance of the proposed scheme, we use it for the benchmark prediction task of learning the Mackey-Glass (MG) chaotic attractor:

$$\frac{dy(t)}{dt} = \frac{\alpha y(t-\tau)}{1 + y(t-\tau)^{\beta}} - \gamma y(t), \tag{5}$$

where $\alpha = 0.2$, $\beta = 10$, $\gamma = 0.1$.

The prior history with a length of N_t is used to predict future data points. Each symbol in the input vector is encoded by the peak amplitude of the Gaussian pulse, as described in the previous section. Figure 3 illustrates nonlinear transformation of the pulse train, corresponding to the input vector x, in SOA. Both input and output powers, shown in the figure, were calculated at $N_c = 25$ redout channels.

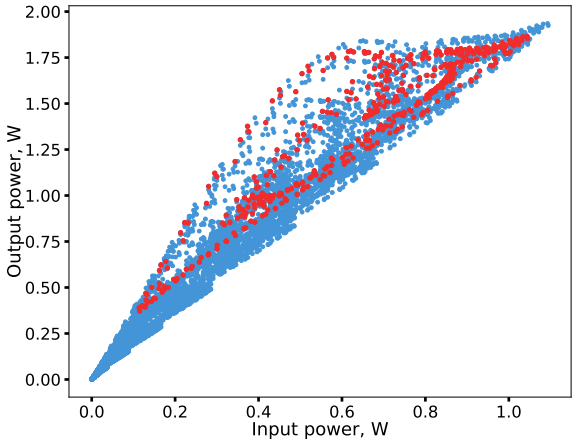

Figure 3. Nonlinear transformation of pulse train instant powers in SOA. $P_0 = 1$ W, $N_t = 512$. Both input and output powers were calculated at $N_c = 25$ redout channels. Red color indicates the peak powers of the input pulses.

We study the learning for one-step ahead prediction task and delay $\tau = 17$. The results are presented in the Table 1. Increasing N_c at the fixed number of the training points $N_t = 256$ leads to growth of the transmission matrix rank and corresponding decrease of the normalized mean squared error $NMSE = \sum_{i=1}^{N}(y_{out}^i - y_t^i)^2/P_0^2$. The optimal number of the output channels N_c is 22 since a further increase in their number does not lead to an increase of the transmission matrix rank. If the number of the training points N_t at the fixed N_c is too small, the matrix rank drops down and the model overfitting takes place (last row in the Table 1).

Nc	**Nt**	**Rank(W)**	**NMSE (Train)**	**NMSE (Test)**
100	256	22	7.4000e-8	1.0808e-7
50	256	22	8.6185e-8	1.1901e-7
25	256	22	1.5379e-7	1.7925e-7
10	256	10	1.0897e-5	1.0316e-5
25	512	20	2.3178e-9	2.4273e-7
25	256	22	1.5379e-7	1.7925e-7
25	128	21	6.0290e-9	1.6923e-7
25	64	3	2.8950e-22	0.0114

Table 1. Training and testing errors for varying number of the redout channels N_c and the training points N_t.

Pulse train average power and duty cycle (T_{FWHM}/T) were also parameters for optimization. Figure 4 shows the testing error dependence on the signal duty cycle at SOA input. The cases of mildly ($\tau = 17$, Fig. 4a) and wildly ($\tau = 30$, Fig. 4b) chaotic attractors were considered. The best prediction results were obtained for duty cycle close to 0.3 in the both considered cases.

Figure 5 demonstrates the Mackey-Glass 50 steps-ahead prediction results for $\tau = 17$. The achieved training NMSE equals to 1.2E-9 and testing NMSE equals to 2.6E-6.

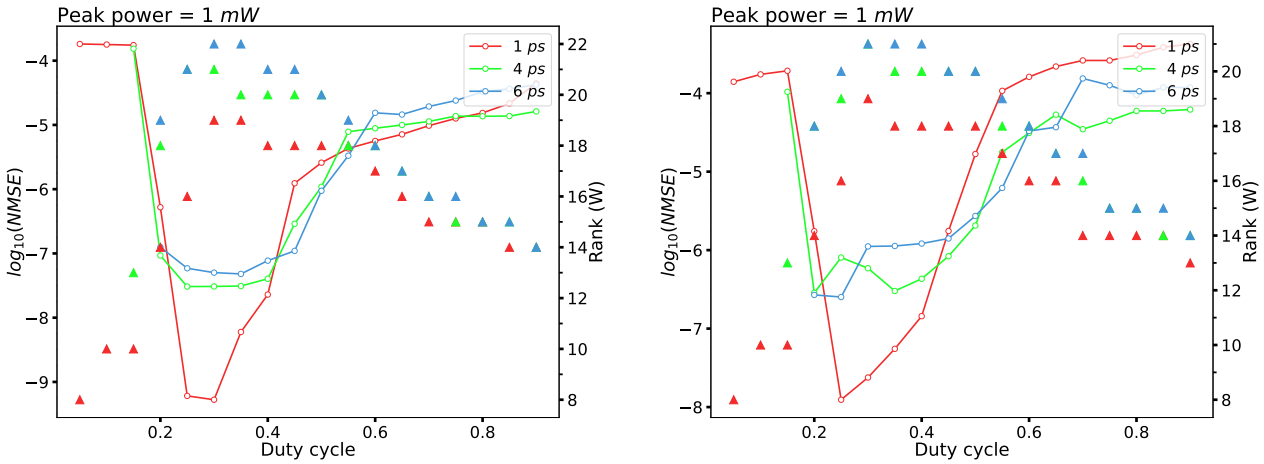

Figure 4. Testing errors for the MG one-step prediction tasks with $\tau = 17$ (a) and $\tau = 30$ (b). $N_t = 512$, $N_c = 25$

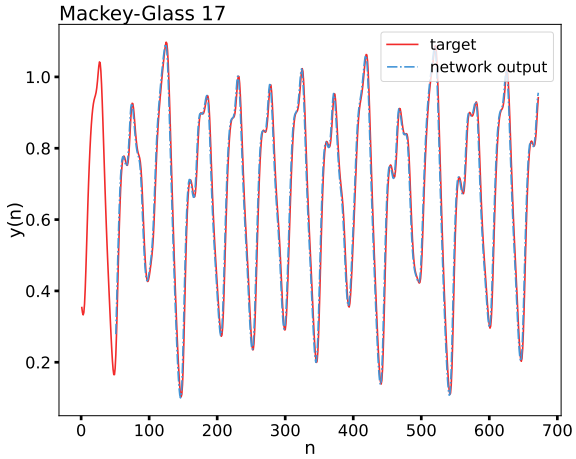

Figure 5. 50-steps ahead MG prediction for $tau = 17$. $N_t = 512$, $N_c = 25$.

ACKNOWLEDGMENTS

The work of D.A.I and A.E.B was supported by the Russian Science Foundation (Grant No.21-42-04401).

REFERENCES

[1] Jaeger, H., "The echo state approach to analysing and training recurrent neural networks-with an erratum note'," *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report* .

[2] Jaeger, H., "Towards a generalized theory comprising digital, neuromorphic and unconventional computing," *Neuromorphic Computing and Engineering* **1**, 012002 (jul 2021).

[3] Sun, C., Song, M., Hong, S., and Li, H., "A review of designs and applications of echo state networks," (2020).

[4] Sorokina, M., Sergeyev, S., and Turitsyn, S., "Fiber echo state network analogue for high-bandwidth dual-quadrature signal processing," *Opt. Express* **27**, 2387–2395 (Feb 2019).

[5] Sorokina, M., "Multidimensional fiber echo state network analogue," *J. Phys. Photonics* (2020).

[6] Marcucci, G., Pierangeli, D., and Conti, C., "Theory of neuromorphic computing by waves: Machine learning by rogue waves, dispersive shocks, and solitons," *Phys. Rev. Lett.* **125**, 093901 (Aug 2020).

[7] Brunner, D., Penkovsky, B., Marquez, B. A., Jacquot, M., Fischer, I., and Larger, L., "Tutorial: Photonic neural networks in delay systems," *Journal of Applied Physics* **124**(15), 152004 (2018).

[8] Vandoorne, K., Van Vaerenbergh, T., Fiers, M., Bienstman, P., Verstraeten, D., Schrauwen, B., and Dambre, J., "Photonic reservoir computing and information processing with coupled semiconductor optical amplifiers," in [*2011 Fifth Rio De La Plata Workshop on Laser Dynamics and Nonlinear Photonics*], 1–3 (2011).

[9] 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).