Optics Letters

Prediction performance of reservoir computing systems based on a diode-pumped erbium-doped microchip laser subject to optical feedback

Romain Modeste Nguimdo,[1,*] Eric Lacot,[2] Olivier Jacquin,[2] Olivier Hugon,[2] Guy Van der Sande,[3] and Hugues Guillet de Chatellus[2]

[1]Optique Nonlinéaire Théorique, Université Libre de Bruxelles, Campus Plaine, CP 231, 1050 Bruxelles, Belgium
[2]Université Joseph Fourier, Grenoble Alpes, CNRS, LIPhy, F-38000 Grenoble, France
[3]Applied Physics Research Group, Vrije Universiteit Brussel (VUB), 1050 Brussels, Belgium
*Corresponding author: Romain. Nguimdo@vub.ac.be

Received 15 November 2016; revised 18 December 2016; accepted 19 December 2016; posted 20 December 2016 (Doc. ID 280955); published 17 January 2017

Reservoir computing (RC) systems are computational tools for information processing that can be fully implemented in optics. Here, we experimentally and numerically show that an optically pumped laser subject to optical delayed feedback can yield similar results to those obtained for electrically pumped lasers. Unlike with previous implementations, the input data are injected at a time interval that is much larger than the time-delay feedback. These data are directly coupled to the feedback light beam. Our results illustrate possible new avenues for RC implementations for prediction tasks.

OCIS codes: (070.4560) Data processing by optical means; (070.6020) Continuous optical signal processing; (140.3490) Lasers, distributed-feedback.

https://doi.org/10.1364/OL.42.000375

Reservoir computing (RC) is a brain-inspired concept for information processing that has been recently demonstrated to be efficient for solving practical time-dependent tasks [1,2]. RC systems operate by ensuring nonlinear mapping between the input and the output, thereby allowing a variety of information processing through training. To perform well, an RC system typically requires high dimensionality and nonlinearity. Traditionally, high dimensionality is obtained by randomly interconnecting a large number of neurons, while the needed nonlinearity can be implemented through sigmoidal activation functions. For example, with an ensemble of 16 interconnected semiconductor optical amplifiers, state-of-the-art performance has been achieved [3].

Alternatively, state-of-the-art performance also has been obtained by relying on a single dynamical nonlinear node subject to delayed feedback [4]. This configuration (typically referred to as the delay-based RC) has the advantage of being easy to train and to implement experimentally. It has led to several implementations, even at high processing speeds, using stand-alone commercial telecommunication components [5–8]. The main differences between these experiments are the type of nonlinearity used and how the input matches with the period of the delay line. In those implementations, the nonlinear response of the reservoir is provided by passive nonlinearity such as saturable absorption of a semiconductor mirror [9–11] or by active devices such as optoelectronic modulators [5,8], optical amplifiers [3], or semiconductor lasers [7]. These experiments have been supported by numerical simulations [8,12–15]. Numerical simulations also have shown that the different modes of multimode lasers subject to optical delayed feedback can be used independently to process independent tasks in parallel [16]. In all cases, the readout layer is trained (using some form of regression) from the state vectors of the reservoir in response to the training data.

Until now, only one experiment has been dedicated to RC systems, in which the processing was done from the response provided by a laser subject to optical delayed feedback [7]. The laser used in this experiment was an electrically pumped single-longitudinal-mode laser, and the input data were either electrically injected by modulating the pump current or optically injected into the reservoir through optical injection using another laser. Note that in other experiments [5,6,8–11], lasers were used as the light source to supply the reservoir.

In this work, we pursue three different objectives. (i) We investigate whether a different type of laser subject to similar delayed feedback can yield similar results to those obtained in [7]. More precisely, we experimentally and numerically investigate whether an optically pumped laser can produce similar results to those of electrically pumped lasers. (ii) In the usual procedure of RC, the injection times of the input data (i.e., the inverse of the processing speeds) are close or correspond to the time delay. In addition, in previous experiments, long time delays (time delays much larger than the system's characteristic time) have been used [4–14,16]. We explore here whether time delays less than or comparable to the laser relaxation time can also be suitable for RC systems. This approach is interesting,

since short delay lines are desirable for on-chip implementations because they consume less wafer space. The time delays will typically be fixed during the manufacturing phase for real-world systems, while the processing speeds can be controlled. We also investigate the effect of such a change for a fixed time delay. (iii) Finally, we explore whether the data can be directly coupled to the feedback light beam instead of using an additional laser for optically injecting the electrical data into the reservoir, as is done in [7].

The RC system can be conceptually divided into three blocks, which are schematically shown in Fig. 1: an input layer, a reservoir, and an output layer. The input layer is the stage where the input data are provided and preprocessed before being injected into the reservoir for their processing. The preprocessing includes the rescaling of the input signal and its multiplication by a mask matrix M that defines the coupling weights from the input to the reservoir. The role of the mask is to ensure the variability of the signal over the different virtual nodes where the information is read out. These virtual nodes are formed by sampling the delay line at a fixed time interval θ. The output layer is the stage where the different node responses are weighted and linearly summed up. The optimal values of these weights are those for which the summation of all the different node responses always approaches the associated target as closely as possible. They can typically be determined with an offline training procedure using digital computers [5–8,11–14] or an online training procedure using a field-programmable gate array [17]. In our case, the training is done offline. Three time scales are relevant: the data injection time T_d, the feedback time-delay T, and the mask length given by $N\theta$, where N is the number of virtual (temporal) nodes. We consider $T_d = N\theta$ and apply this same duration during the postprocessing of each data point.

As the benchmark task, we will use the Santa Fe time series to evaluate the prediction performance of the system. The Santa Fe data are an intensity time series experimentally recorded from a far-IR laser operating in a chaotic state [18]. The goal for this task is to predict the sample one step ahead in a chaotic time trace before it has been injected into the reservoir computer. The system performance is evaluated by calculating the normalized mean square error (NMSE) between the predicted value y and the expected value $\bar{y}$:

$$\mathbf{NMSE}(\mathbf{y}, \bar{\mathbf{y}}) = \frac{\langle ||y(n) - \bar{y}(n)||^2 \rangle}{\langle ||\bar{y}(n) - \bar{y}(n)\rangle \bar{y}(n)\langle ||^2 \rangle}, \qquad \textbf{(1)}$$

where n is a discrete time index and $\|.\|$ and $\langle . \rangle$ stand for the norm and the average over time, respectively. Note that NMSE = 0 represents a perfect prediction, while NMSE = 1 indicates no prediction at all. Our results will be compared to those in [7], as it is a similar RC system, which was, however, built using a different type of laser and a large time delay.

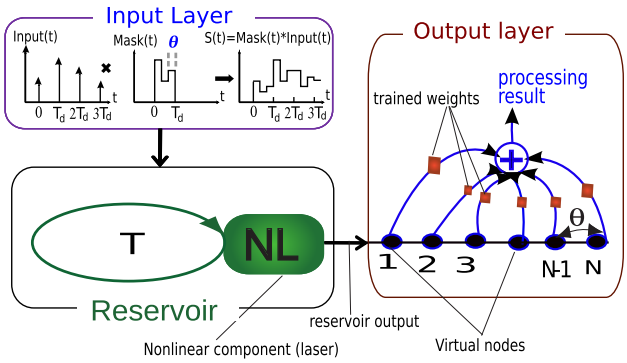

Fig. 1. Conceptual scheme of RC based on a nonlinear node (laser) with delayed feedback.

The experimental setup is shown in Fig. 2. It is composed of a diode-pumped erbium-doped microchip laser with optical delayed feedback. The microchip laser is a sample provided by the LETI-CEA laboratory (Technologies Avanckes) [19]. The threshold currents of the diode pump and the microlaser are $I_{sp} \approx 24.2$ mA and $I_s \approx 100$ mA, respectively. The microchip laser's stability threshold is close to, but above, its lasing threshold. The microlaser is bi-mode, which, above I_s, emits a total radiation of a few milliwatts at the wavelengths $\lambda_1 \sim 1532$ nm and $\lambda_2 \sim 1535$ nm, with a side mode suppression of ≈26 dB. The emitted light has been collected using collimating lensed fiber with a collection efficiency of ≈50%. To implement the delayed feedback, the collected light beam is split into two parts using a 50% optical coupler: one part is directly read out by a photodiode with 1 GHz bandwidth, while the second part is connected to port 2 of an optical circular. The circulator couples light from port 2 to port 3 and from port 3 to port 1 with minimal power loss. The signal at port 3 is delayed by a single-mode optical fiber of length ≈126 m (i.e., time delay of ∼630 ns). The light beam at the fiber output is split into two paths. One path passes through a 10 GHz $LiNbO_3$ phase modulator (PM) (half-wave voltage $V_\pi = 5$ V, maximum input power 28 dBm). The PM output signal is combined with the light beam of the second path using a 2 × 2 optical coupler implementing an intensity modulation of the light beam. One output port of this coupler is connected to port 1 of the circulator, while the second output port is connected to an optical spectrum analyzer. Thus, taking into account the losses, about 10% of the collected power is reinjected into the laser.

The original data to be processed are first convoluted with a random mask that has four discrete values (−1, −0.25, 0.25, 1) generated randomly with equal probability at a time interval of $\theta = 24$ ns. The resulting signal is uploaded in a two-channel arbitrary waveform generator (AWG) (Tektronix AFG3102C, dual channel, 1 GS/s, 100 MHz), from which it is injected into the reservoir via the PM rf electrode. The amplitude of the injected signal is rescaled to obtain an optimal contrast of the modulation at the laser output. In our case, the peak-to-peak voltage from the AWG was 10 V. The data are recorded using a digital oscilloscope (LeCroy, 200 MHz, 2.5 GS/s). The AWG is perfectly synchronized with the oscilloscope so that the original data and the laser response to these data can be simultaneously recorded. We experimentally use 1000 steps for training and

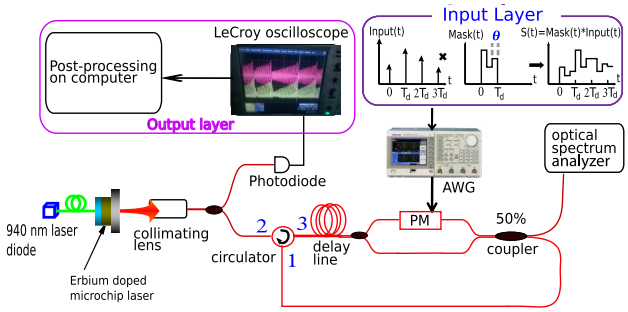

Fig. 2. Experimental setup. PM, phase modulator; AWG, arbitrary waveform generator.

250 steps for testing. The samples are injected at the time interval $T_d = 2.4$ μs, which is 3.8 times larger than the time delay, i.e., $T = 630$ ns. This corresponds to a reservoir with $N = T_d/\theta = 100$ virtual nodes. As such, only $\sim N/3.8$ virtual nodes receive an input corresponding to a previous time.

Figure 3 shows the temporal profiles of the original data (black) and the laser response at the photodiode output (red) for two values of the diode-pumped currents with one close to the threshold ($I = 120$ mA) and another far away from the threshold ($I = 220$ mA). For $I = 120$ mA, the microlaser emits a low power, and therefore the fraction of the signal (i.e., about 10% of the collected power) to be reinjected into the laser is small. In this case, the dynamical response of the laser to the input signal may also be limited, since the feedback light beam is the one carrying the input data. For $I = 220$ mA, the beam fraction driving the data is larger. The laser response is quite similar to the input signal. These laser output signals are used to optimize the readout weights. The prediction error rate NMSEs obtained for the two cases were ≈0.4 and ≈0.07 for $I = 120$ mA and $I = 220$ mA, respectively. To demonstrate that the nonlinearity of the laser is necessary for this computation, we have also optimized the weights directly from the AWG's output signals. We obtained a prediction error rate NMSE of ≈0.7. This significantly worse performance shows that the eventual nonlinearity caused by out-of-bandwidth operation of the AWG does not already allow for reservoir computing.

To identify the most suitable parameter regimes for which the system can successfully predict a chaotic input signal one time step ahead in the future, we show in Fig. 4 the system performance expressed by the NMSE as a function of the diode pump rate $\mu = (I - I_{sp})/(I_s - I_{sp})$. For each value of μ, we run the experiments six different times and record for each run a time trace of 5 ms at the output of the photodiode. The sampling rate at the oscilloscope is 2 ns. The NMSE values shown are the mean values over the runs. For low pump rates μ, large NMSE values are obtained, indicating worse performance. This degradation is due to noise, which is dominant for small laser output powers [the signal-to-noise ratio (SNR) is small in this case]. This noise is mainly readout noise from the photodiode. For large values of μ (i.e., large laser output signals), the SNR is large and the influence of the readout noise can be minimized. For such cases, we obtain the lowest NMSE of 0.12 ± 0.04 at $\mu \approx 2.58$ (note that for clarity, the error bars are not plotted), while a lowest NMSE of 0.108 was obtained in [7] for the

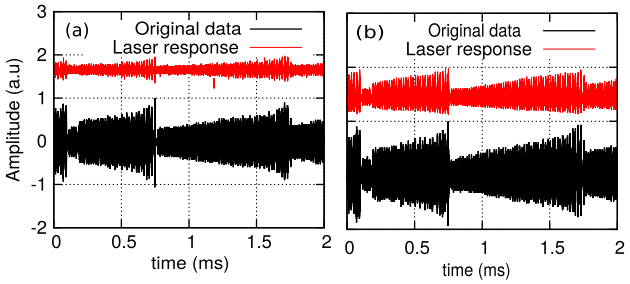

Fig. 3. Original data (black) and the corresponding experimental recorded signal of the laser response read by the photodiode (red) for the diode pump current (a) $I = 120$ mA and (b) $I = 220$ mA. The amplitudes of the input data and the laser response have been rescaled with the maximum value in each case. The laser response signal has been shifted for clarity.

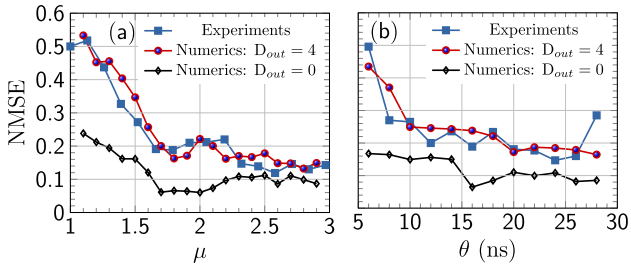

Fig. 4. Prediction error NMSE as a function of (a) the pump rate μ for a virtual node separation of $\theta = 24$ ns, (b) θ for $\mu = 2.5$.

same task. The performance of the two systems is therefore similar, although we consider a time delay comparable to the laser's relaxation oscillation time τ_{RO} [note that $\tau_{RO} \approx (\gamma\gamma_e(\mu - 1))^{-1/2} \approx 750$ ns at $\mu \approx 2.5$].

For real-world applications, the delay length would be fixed during the manufacturing phase of the device, while the length of the mask could be independently varied by changing, for example, the temporal node's separation. To demonstrate this fact, we show in Fig. 4(b) the NMSE values for different values of θ when N and μ are kept fixed to 100 and 2.5, respectively. The optimal value lies around $\theta = 24$ ns. These are therefore the values of θ for which the optimal coupling between the virtual nodes is obtained.

For further insights, we performed numerical simulations using the model of the diode-pumped erbium microchip laser subject to optical delayed feedback introduced in [20,21]. The relavant variables of the model are the slow-varying complex envelope of the electric field $E_m(t)$ associated with the longitudinal mode m (with $m = 1, 2$) and its corresponding carrier number $N_m(t)$:

$$\dot{E}_{1,2} = \frac{1}{2}(Bg_{1,2}N_{1,2} - \gamma)E_{1,2} + \frac{i\alpha}{2}BN_{1,2} + \frac{\eta}{2}[1 + e^{i(S(t)-\omega_{1,2}\delta T)}]E_{1,2}(t - T)e^{i\omega_{1,2}T} + \xi_{1,2}(t),$$

$$\dot{N}_{1,2} = \gamma_e(N_p - N_{1,2}) + gBN_{1,2}(|E_{1,2}|^2 + \rho|E_{2,1}|^2) + \zeta_{1,2}(t),$$

where the parameters are the Einstein coefficient B, the decay rate of the population inversion γ_e, the pumping rate $\gamma_e N_p$, the laser cavity decay rate γ, the linewidth enhancement factor α, the feedback rate η, the time delay T, the cross-saturation parameter ρ, and the solitary laser frequency for mode m, ω_m. $S(t)$ is the signal that results from the convolution between the original data to be processed and the mask. The noise is modeled as the Langevin forces $\xi_{1,2}(t)$ and $\zeta_{1,2}(t)$, which describe the quantum fluctuations of the laser population and the radiation field. These forces are defined as having a zero mean value and white-noise-type correlation functions $\langle \xi_i(t)\xi_j^*(t')\rangle = D\delta_{ij}(t - t')$ and $\langle \zeta_i(t)\zeta_j^*(t')\rangle = D\gamma\gamma_e\mu\delta_{ij}(t - t')/B$ for $i, j = 1, 2$, where D is the spontaneous emission factor.

From the experiments, we have retrieved the parameters $\gamma = 53.66$ μs^{-1}, $\gamma_e = 0.217$ μs^{-1}, and $\rho = 0.43$. We also consider other parameters from [20,21]: $B = 122.1$, $\alpha = 1$, $\eta = 0.3$ μs^{-1}, $g = 0.95$, $N_p = \mu\gamma/B$, $\omega_m T = 0$, $\omega_m \delta T = -0.9\pi$, and $D = 10^{-4}$. We numerically use 3000 steps in the data set for training and 1000 other steps for testing. For each value of μ, we run the simulations 10 different times and record for each run the laser output as $|E(t)|^2 = |E_1(t) + E_2(t)|^2$.

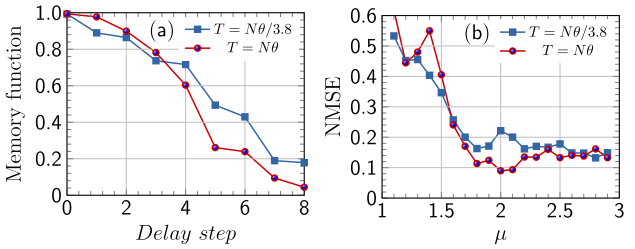

Fig. 5. Numerical results. (a) Memory function for two values of time delay. For $T = N\theta$, all N virtual nodes are connected to nodes in the previous state. For $T = N\theta/3.8$, only $N/3.8$ virtual nodes are connected to nodes in the previous state. (b) Corresponding NMSE as a function of the pump rate μ. We consider $\theta = 24$ ns and $\mu = 2.5$.

Additional Gaussian white noise is added to $|E(t)|^2$ to model the readout noise from the photodiode. Thus, the weights are optimized from $|E(t)|^2 + D_{\text{out}}\xi_{\text{out}}(t)$ instead of $|E(t)|^2$, where D_{out} is the noise amplitude and $\xi_{\text{out}}(t)$ is the Gaussian white noise with zero mean and correlation $\langle \xi_{\text{out}}(t)\xi^*_{\text{out}}(t')\rangle = \delta(t - t')$. $D_{\text{out}} = 0$ refers to noiseless photodetectors. We choose $D_{\text{out}} = 4$, as it is the value for which numerical and experimental SNRs, calculated as $\text{SNR} = 10\log_{10}(\langle |E(t)|^2\rangle / \langle D_{\text{out}}\rangle)$, are similar. The SNR is ≈10 dB at $\mu = 1.2$ and ≈26 dB at $\mu = 2.5$. The value of D_{out} is kept fixed for the rest of the work in this Letter.

The numerical obtained NMSE values as a function of the pump rates are also shown in Fig. 4 in comparison with the experimental values. Again, the error bars are not plotted. In both cases, excellent agreement is found between experimental and numerical results. In particular, it is confirmed that because of noise, the system performance strongly degrades for small values of μ [see Fig. 4(a)], and also for high processing speeds [see Fig. 4(b)]. We obtain, for example, NMSE ≈ 0.43 for $\theta = 6$ ns, although all the N virtual nodes are connected through the feedback to nodes in the previous state for this θ. Note that for fixed N, the processing speed increases with decreasing θ. By comparing the NMSE values obtained from a system with (circles) and without readout noise (diamonds), we find that the degradation of the NMSE for low values of μ is mainly due to the readout noise. Through numerical simulations, we have also found that calculating the weights from $|E_1(t)|^2$ or $|E_2(t)|^2$ (i.e., considering only one mode) yields similar results to calculations using $|E_1(t) + E_2(t)|^2$. This is not surprising, since the mode spacing is large, such that the heterodyne signal is outside the detection bandwidth of the system. In this case, the choice of a monomode or bimode laser is not important.

The memory capacity of the system may decrease when only a fraction of the virtual nodes are connected through the feedback to the previous input states. To demonstrate this fact, we compare in Fig. 5(a) the simulated results of the memory function (calculated from a uniformly distributed random signal drawn in the interval $[-0.5; 0.5]$) when $N/3.8$ virtual nodes (squares) and when all N virtual nodes (circles) are connected to nodes in a previous state considering $\theta = 24$ ns and $\mu = 2.5$. Effectively, it is seen that the memory slightly decreases for $N/3.8$ virtual nodes connected to nodes in a previous state. The decreasing slope of the memory is also slow in this case. However, the calculation of the NMSE shows that these changes in the memory do not significantly degrade the system performance for one-step-ahead prediction [Fig. 5(b)]. However, the low memory in Fig. 5(a) suggests that this system may not be suitable for some tasks requiring large memory capacity.

In conclusion, we have experimentally and numerically shown that diode-pumped erbium-doped microchip lasers subject to optical feedback can be used to implement RC systems for prediction tasks. Using a Santa Fe time series as a benchmark, we found a best prediction error similar to that obtained with a similar system having a long time delay [7], but our system used of a short time delay. Our results also have shown that even multimode lasers with large mode spacing can be used.

Funding. Fonds De La Recherche Scientifique (FNRS); Belgium Science Policy Office (IAP-7/35 photonics@be).

Acknowledgment. R. M. N. acknowledges the support of CNRS during his stay at LIPhy.

REFERENCES

1. W. Maass, T. Natschläger, and H. Markram, Neural Comput. **14**, 2531 (2002).
2. H. Jaeger and H. Haas, Science **304**, 78 (2004).
3. K. Vandoorne, P. Mechet, T. V. Vaerenbergh, M. Fiers, G. Morthier, D. Verstraeten, B. Schrauwen, J. Dambre, and P. Bienstman, Nat. Commun. **5**, 4541 (2014).
4. L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, Nat. Commun. **2**, 468 (2011).
5. L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutierrez, L. Pesquera, C. R. Mirasso, and I. Fischer, Opt. Express **20**, 3241 (2012).
6. Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, and S. Massar, Sci. Rep. **2**, 287 (2012).
7. D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, Nat. Commun. **4**, 1364 (2013).
8. M. C. Soriano, S. Ortín, D. Brunner, L. Larger, C. R. Mirasso, I. Fischer, and L. Pesquera, Opt. Express **21**, 12 (2013).
9. F. Duport, B. Schneider, A. Smerieri, M. Haelterman, and S. Massar, Opt. Express **20**, 22783 (2012).
10. A. Dejonckheere, F. Duport, A. Smerieri, L. Fang, J.-L. Oudar, M. Haelterman, and S. Massar, Opt. Express **22**, 10868 (2014).
11. Q. Vinckier, F. Duport, A. Smerieri, K. Vandoorne, P. Bienstman, M. Haelterman, and S. Massar, Optica **2**, 438 (2015).
12. K. Hicke, M. A. Escalona-Moran, D. Brunner, M. C. Soriano, I. Fischer, and C. R. Mirasso, IEEE J. Sel. Top. Quantum Electron. **19**, 1501610 (2013).
13. R. M. Nguimdo, G. Verschaffelt, J. Danckaert, and G. Van der Sande, Opt. Express **22**, 8672 (2014).
14. J. Nakayama, K. Kanno, and A. Uchida, Opt. Express **24**, 8679 (2016).
15. R. M. Nguimdo, G. Verschaffelt, J. Danckaert, and G. Van der Sande, Opt. Express **24**, 1238 (2016).
16. R. M. Nguimdo, G. Verschaffelt, J. Danckaert, and G. Van der Sande, IEEE Trans. Neural Netw. Learn. Syst. **26**, 3301 (2015).
17. P. Antonik, F. Duport, A. Smerieri, M. Hermans, M. Haelterman, and S. Massar, "Online training of an opto-electronic reservoir computer," in *Proceedings of International Conference on Neural Information Processing, Part II* (2015), pp. 233–240.
18. A. S. Weigend and N. A. Gershenfeld, "Time series prediction: forecasting the future and understanding the past," 1993, ftp://ftp.santafe.edu/pub/Time-Series/Competition.
19. P. Thony and E. Molva, *OSA Trends in Optics and Photonics on Advanced Solid-State Lasers*, S. A. Payne and C. Pollock, eds. (Optical Society of America, 1996), Vol. **1**, paper IL3.
20. E. Lacot, O. Jacquin, O. Hugon, and H. Guillet de Chatellus, Appl. Opt. **54**, 9763 (2015).
21. C. Szwaj, E. Lacot, and O. Jacquin, Phys. Rev. A **70**, 033809 (2004).