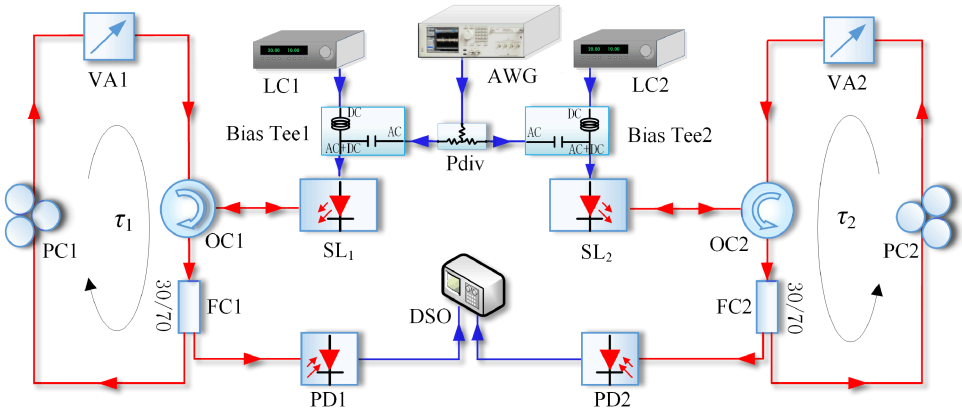
IEEE Photonics Journal
An IEEE Photonics Society Publication

Experimental Investigation of an Optical Reservoir Computing System Based on Two Parallel Time-Delay Reservoirs

Volume 13, Number 3, June 2021

Dian-Zuo Yue
Yu-Shuang Hou
Zheng-Mao Wu
Chun-Xia Hu
Zhen-Zhen Xiao
Guang-Qiong Xia



DOI: 10.1109/JPHOT.2021.3075055

Experimental Investigation of an Optical Reservoir Computing System Based on Two Parallel Time-Delay Reservoirs

Dian-Zuo Yue,[1] **Yu-Shuang Hou**,[2] **Zheng-Mao Wu**,[1] **Chun-Xia Hu**,[1,3] **Zhen-Zhen Xiao**,[1] **and Guang-Qiong Xia**[1]

[1]School of Physical Science and Technology, Southwest University, Chongqing 400715, China
[2]School of Science, Inner Mongolia University of Science and Technology, Baotou 014010, China
[3]College of Mobile Telecommunications, Chongqing University of Posts and Telecom, Chongqing 401520, China

DOI:10.1109/JPHOT.2021.3075055

Manuscript received March 3, 2021; revised April 16, 2021; accepted April 20, 2021. Date of publication April 22, 2021; date of current version May 3, 2021. This work was supported in part by the National Natural Science Foundation of China under Grants 61775184, 61875167, and 62065015, in part by the Postgraduate Research and Innovation Project of Chongqing Municipality under Grant CYB19087, and in part by the Natural Science Foundation of Inner Mongolia Autonomous Region of China under Grant 2019MS06022. Corresponding authors: Zheng-Mao Wu and Guang-Qiong Xia (e-mail: zmwu@swu.edu.cn; gqxia@swu.edu.cn).

Abstract: We experimentally investigate the performances of an optical reservoir computing (RC) system based on two parallel time-delay reservoirs composed of two semiconductor lasers (SLs) subject to optical feedback. In such a system, the information being processed is split into two parts to send into two reservoirs through directly modulating the pump currents of two SLs, and the temporal output of the two SLs are sampled and taken as the virtual node states for training and testing. Via Santa Fe time series prediction task and multi-waveform recognition task, the performances of the proposed RC system are investigated and compared with those of the system based on one reservoir. The results show that the system based on two parallel reservoirs behaves better performance and stronger parameter robustness than that based on one reservoir. Moreover, through analyzing the dependence of the system performances on the number of virtual node states actually used for readout, the potential data processing rate (DPR) of the system is evaluated. For processing a prediction task under guaranteeing the normalized mean square error below 0.1 and a recognition task under guaranteeing the signal error rate below 0.005, the potential DPR of the proposed RC system can achieve 200 MSa/s, which is twice the DPR of the system with only one reservoir.

Index Terms: Reservoir computing (RC), semiconductor laser (SL), time-delay reservoirs, chaotic time series prediction, multi-waveform recognition.

1. Introduction

Nowadays, with the advent of big data era, the demand for efficient information processing technique is growing rapidly. Besides traditional algorithm of Von Neumann machines, artificial neural networks (ANNs) inspired by human brains have shown excellent performance in dealing with complex tasks such as chaotic time series prediction and pattern recognition [1].

Among ANNs, recurrent neural network (RNN) is capable of performing time-dependent tasks, but it is time-consuming to train the network weights of the input layer, the hidden layer, and the output layer. In order to overcome this drawback in RNN, the concept of reservoir computing (RC) is proposed [2], [3]. RC possesses a simplified training procedure since only the output weights need to be trained [4], [5]. Moreover, because the input weights and the internal connection weights of reservoir are randomly generated and fixed, RC is more suitable for hardware implementation than RNN. Traditionally, the reservoir is composed of a large number of nonlinear nodes, and such an RC is usually named as spatially distributed RC [6]. It has been demonstrated that the reservoir in RC can also be composed of a single nonlinear node with a time-delay feedback loop, and the corresponding RC is named as time-delay RC (TDRC) [7]. Comparatively speaking, TDRC is easier to implement than spatially distributed RC, since its reservoir consists of only a single nonlinear node and a feedback loop. In recent years, several groups have implemented TDRCs based on different time-delay systems including electronics [7], [8], optoelectronics [9]–[13], and optics [14]–[17].

TDRC facilitates hardware implementation but pays a price in terms of data processing rate (DPR) owing to its serial feeding and reading procedure. For TDRC, the original data to be processed is first expanded over a time T and multiplied by a mask [18], [19]. After masking procedure, the resulted sequence in one period of T is serially fed into a reservoir, and the reservoir states (called virtual node states) are serially obtained by sampling the state of the single nonlinear node n times in a period T with an interval θ. As a result, the DPR of TDRC ($=T^{-1}$) is determined by the number of virtual nodes n and the interval θ between virtual nodes. Usually, n is set between 10^2 and 10^3 to ensure a high dimensional state space, and θ is governed by the response time of the nonlinear node as well as the sampling rate of a readout device [20]. The setting of n and θ in different TDRC systems result in obvious differences in DPR. For instance, the DPR of optoelectronic TDRCs based on a Mach-Zehnder modulator implemented by Larger *et al.* [9] and by Paquot *et al.* [10] are about 48 KSa/s ($n = 400$, $\theta \approx 52.18\,ns$) and 117 KSa/s ($n = 50$, $\theta \approx 170$ *ns*), respectively. The DPR of an all-optical TDRC based on a semiconductor optical amplifier implemented by Duport *et al.* is about 126 KSa/s ($n = 50$, $\theta = 155.76$ *ns*) [14]. In recent years, a type of all-optical TDRC based on semiconductor laser (SL) has shown relatively high DPR due to the fast response characteristic of SLs. The DPRs of SL-based TDRC implemented by Brunner *et al.* [15] and Kuriki *et al.* [21] reach to 13 MSa/s ($n = 388$, $\theta = 0.2\,ns$) and 28 MSa/s ($n = 177$, $\theta = 0.2\,ns$), respectively. Vatin *et al.* successfully processed two tasks with a DPR of 51.3MSa/s ($n = 492$, $\theta = 0.08\,ns$) through a vertical-cavity surface-emitting laser based TDRC [16]. Based on a photonic integrated SL chip, Takano *et al.* [22] and Harkhoe *et al.* [23] achieve high DPRs of 0.806 GSa/s ($n = 124$, $\theta = 0.01\,ns$) and 0.87 GSa/s ($n = 23$, $\theta = 0.05\,ns$), respectively, which represent the state-of-the-art level of DPR for TDRC. However, further increasing the DPR of a TDRC system is very challenging, because for achieving such high DPRs, the system not only requires to further reduce the number of virtual nodes, but also requires a readout device (usually is a digital storage oscilloscope) with a sampling rate higher than tens of GSa/s.

In view of the DPR limitation of TDRC, another method to implement RC that using ensemble of time-delay reservoirs was proposed by Ortín *et al.* [24], which combines the merits of high DPR of spatially distributed RC and easy implementation of TDRC. In theory, by collecting virtual node states simultaneously from multiple time-delay reservoirs, information processing time can be reduced [25]. In recent years, although RC based on multiple time-delay reservoirs have been numerically investigated in either coupled or uncoupled time-delay systems [26]–[30], the relevant experimental investigation is still lacking. In our previous work [31], we numerically investigated the performance of an RC system based on two parallel reservoirs, which shown higher DPR and easier to implement in hardware. Therefore, it is valuable to perform experimental research on the basis of simulation work to further verify the feasibility of the scheme.

In this work, we experimentally demonstrate an optical RC system based on two parallel time-delay reservoirs, where the two reservoirs are composed of two SLs subject to optical feedback, respectively. Via an arbitrary waveform generator (AWG), two identical information are sent into

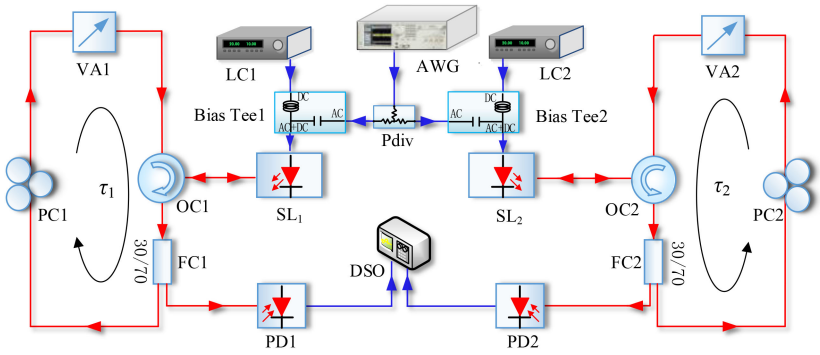


Fig. 1. Experimental setup of RC based on two parallel time-delay reservoirs. FC: Fiber coupler; VA: Variable attenuator; PC: Polarization controller; OC: Optical circular; PD: Photo detector; LC: Laser controller; Pdiv: Power divider; AWG: Arbitrary waveform generator; DSO: Digital storage oscilloscope.

two reservoirs through directly modulating the pump current of two SLs, respectively. The output intensities of two SLs are sampled as virtual node states for training and testing. A chaotic time series prediction task and a multi-waveform recognition task are employed to test the prediction performance and classification performance of the system, respectively. Under different feedback ratio and bias current of two SLs, the performances of the system are examined and compared to the system with a single reservoir. Moreover, through analyzing the dependence of the system performance on the number of the selected virtual node states in readout layer, we evaluate the potential DPR that this system can achieve for processing the tasks under maintaining good performances.

2. Experimental Setup

Fig. 1 shows the experimental setup of the RC system based on two time-delay reservoirs. Two distributed feedback SLs (SL_1 and SL_2) subject to optical feedback are utilized as two parallel reservoirs. Two laser controllers (LCs) are used to control two SLs, where the current and temperature control accuracy of LCs are 0.01 mA and 0.01 °C, respectively. Throughout the experiment, the temperatures of two SLs are all set at 20.00 °C, and the threshold currents of SL_1 and SL_2 are about 8.3 mA and 8.6 mA, respectively. A single channel AWG (Tektronix, AWG70001A, 1.5 KSa/s-50 GSa/s) is used to output the masked information with a peak to peak voltage (V_{pp}) of 500 mV. Via a power divider (Picosecond, 5331), the masked information is split into two parts, which are injected into RF ports of two Bias Tees (Picosecond, 5541A, 80 KHz-26 GHz) to modulate the pump current of two SLs, respectively. The optical feedback loop of each SL is consisted of an optical circular (OC), a 30:70 fiber coupler (FC), a variable attenuator (VA), and a polarization controller (PC). VA1 (VA2) is utilized to adjust the feedback ratio k_1 (k_2) of the feedback loop SL_1 (feedback loop SL_2), where k_1 (k_2) is defined as the ratio between the feedback intensity and output intensity of SL_1 (SL_2). The delay times of feedback loop SL_1 and SL_2 are $\tau_1 = 30.7$ *ns* and $\tau_2 = 31$ *ns*, respectively. The 30% part of the FC1 (FC2) is converted into electrical signal by the photo detector (New Forus, 1544-B, 12 GHz bandwidth) PD1 (PD2), and then the electrical signal is recorded by a digital storage oscilloscope (DSO, Agilent, DSO-X 91604A, 40 GSa/s, 16 GHz bandwidth) for off-line training and testing.

Fig. 2 shows the schematic illustration of the RC based on two parallel time-delay reservoirs. The experimental system can be divided into three parts: an input layer, a reservoir, and an output layer. At the input layer, the original data undergoes a preprocessing procedure named as the masking. Firstly, the original data is sampled and held for a period T to generate an input stream $u(t)$. Secondly, $u(t)$ is multiplied by a mask signal to form the input information $s(t)$, where the

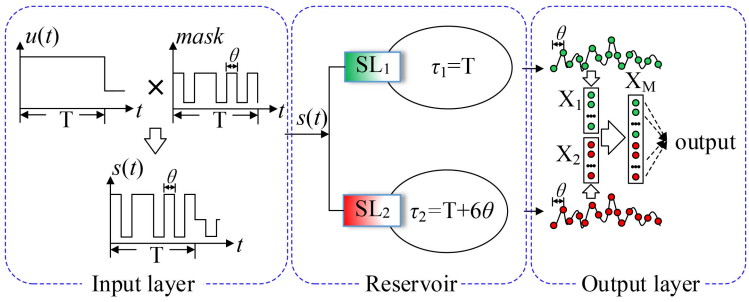


Fig. 2. Schematic illustration of the RC based on two parallel time-delay reservoirs.

mask signal used in this work is a binary sequence randomly extracted from {0.1, 1}. Finally, the information $s(t)$ is divided into two parts and sent to two parallel time-delay reservoirs, respectively. In this work, we set $T = \tau_1 = 30.7$ *ns* and $\theta = 0.05$ *ns*. Within a period T, the information $s(t)$ is a piecewise constant over 614 intervals of θ, and $s(t) = \boldsymbol{s}(n)$ when $(n-1)\theta \leq t < n\theta$, for $n = 1, 2, \ldots, 614$. Therefore, for the reservoir of SL_1, delay time τ_1 equals to the input period T, which corresponds to a synchronized regime [15]. For the reservoir of SL_2, delay time τ_2 (31 *ns*) equals to $T+6\theta$, which corresponds to an unsynchronized regime [10]. Under electrical information modulation and optical feedback, the two SLs will present complex nonlinear transient responses which map low-dimensional input data to a state space with higher dimensionality [32], [33]. In the readout layer, the virtual node states are obtained by sampling the temporal output of two SLs with an interval θ. Consequently, within a period T, two state vectors containing 614 virtual node states are collected from two reservoirs, respectively. For instance, within the i-th period T, the input information is $\boldsymbol{s}^i(n) = [s^i(1), s^i(2), \ldots, s^i(614)]$. For the reservoir of SL_1, the state vector is denoted as:

$$\boldsymbol{x}_1^i(n) = f_1(s^i(n) + k_1 x_1^{i-1}(n)), \quad 1 \leq n \leq 614 \tag{1}$$

For the reservoir of SL_2, under unsynchronized regime, 620 states containing previous input information are remained in the feedback loop. When $\boldsymbol{s}^i(n)$ is injected, we collect 614 states of SL_2 corresponding to the currently input of $\boldsymbol{s}^i(n)$, where the state vector can be expressed as:

$$\boldsymbol{x}_2^i(n) = \begin{cases} f_2(s^i(n) + k_2 x_2^{i-1}(n-6)), & 7 \leq n \leq 614 \\ f_2(s^i(n) + k_2 x_2^{i-2}(614-6+n)), & 1 \leq n \leq 6 \end{cases} \tag{2}$$

Where f_1 and f_2 in (1) and (2) denotes the nonlinear transformation of SL_1 and SL_2, respectively. The two state vectors of $\boldsymbol{x}_1^i(n) = [x_1^i(1), x_1^i(2), \ldots, x_1^i(614)]^{\mathrm{T}}$ and $\boldsymbol{x}_2^i(n) = [x_2^i(1), x_2^i(2), \ldots, x_2^i(614)]^{\mathrm{T}}$ constitute the i-th column of state matrices X_1 and X_2, respectively. Meanwhile, $\boldsymbol{x}_1{}^i(n)$ and $\boldsymbol{x}_2{}^i(n)$ can be merged as a new state vector as $\boldsymbol{x}_{\mathrm{M}}^i(n) = [x_1^i(1), x_1^i(2), \ldots, x_1^i(614), x_2^i(1), x_2^i(2), \ldots, x_2^i(614)]^{\mathrm{T}}$ which constitute the i-th column of state matrix X_{M}. After inputting L original data, two state matrixes X_1 and X_2 of dimension $614 \times L$ and the one merged state matrix X_{M} of dimension $1228 \times L$ can be obtained. In this work, based on the state matrix X_{M} we perform RC (named as RC_{M}). Meanwhile, for comparison, X_1 (X_2) is also used independently for performing RC_1 (RC_2) as that in a system based on a single reservoir. In the training procedure, the output weights are optimized using a ridge regression method [15], which is expressed as:

$$W_{out} = (X^{\mathrm{T}}X + CI)^{-1}X^{\mathrm{T}}y_d \tag{3}$$

where X denotes the state matrix used for readout, y_{d} is the desired value, $C \in R^+$, and I is the identity matrix. In order to avoid over-fitting, we set $C = 5 \times 10^{-3}$ in this work.

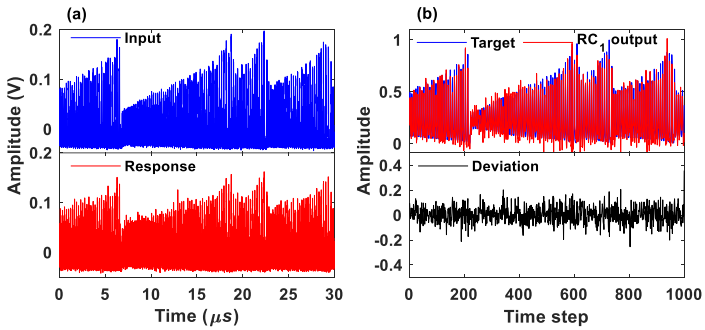


Fig. 3. Temporal waveforms of Santa Fe time series prediction task. (a) Input and response of SL_1, (b) prediction results of RC_1.

3. Results and Discussion

3.1 Santa Fe Time Series Prediction Task

In this section, we perform Santa Fe time series prediction task to evaluate the prediction performance of the system. Santa Fe time series containing 9000 data points was generated by a far-infrared laser operating in a chaotic regime [34]. In this test, we use the front 3000 points of the data for training and the next 1000 points for testing. The target function of this task is the Santa Fe time series but one step ahead of the input data. We use normalized mean square error (NMSE) to quantify the prediction performance, and the NMSE is defined as:

$$NMSE = \frac{1}{L}\sum_{n=1}^{L}(y_{out}(n) - y_d(n))^2/var(y_d) \tag{4}$$

where L is the number of data used for testing, y_d denotes desired value, y_{out} denotes RC output, and *var* represents the variance. We consider the system possesses a good prediction performance when NMSE is lower than 0.1.

Fig. 3 shows an example of temporal waveforms for this task, which is obtained from RC_1 under bias current $j_1 = 10$ mA and feedback ratio $k_1 = -30$ dB, respectively. The upper panel in Fig. 3(a) shows the masked information injected into SL_1, and the lower panel in Fig. 3(a) shows the response of SL_1 under electrical information modulation and optical feedback. Prediction results of RC_1 are shown in the upper panel of Fig. 3(b), which is similar to the target value. Meanwhile, the deviation between RC_1 output and target is shown in the lower panel of Fig. 3(b), where the NMSE is 0.095.

In order to systematically study the performance of RC_M based on two reservoirs and compare its performance with that of RC_1 and RC_2, we next perform RC_1, RC_2, and RC_M under different parameters of SL_1 and SL_2. Note that, since the two SLs have a similar power-current characteristic, in the following tests, we change the parameters of two SLs simultaneously in a same range. The NMSEs as a function of feedback ratio (k_1, k_2) and bias current (j_1, j_2) are plotted in Fig. 4(a) and Fig. 4(b), respectively. In Fig. 4(a), we set $j_1 = j_2 = 10$ mA and simultaneously increase k_1 and k_2 from -40 dB to -10 dB. As shown in this diagram, the NMSE curves for RC_1, RC_2, and RC_M first decrease and then gradually increase with the increasing of feedback ratio, where the minimum values for three curves are achieved at $k_1 = k_2 = -35$ dB. The variation trend of NMSEs may indicates that a too small feedback ratio cannot provide sufficient memory for this task, and a too large feedback ratio causes system oscillation thereby reduces consistency property [15, 35]. We can find that when the feedback ratio is less than -27 dB, RC_1 and RC_2 can perform well (NMSE < 0.1), and the minimum NMSEs for RC_1 and RC_2 are 0.06 and 0.05, respectively. Based on the merged virtual node state matrix X_M, the NMSE for RC_M is always smaller than that for RC_1 and RC_2. Moreover, the NMSE for RC_M is less than 0.1 until the feedback ratio is larger than -15 dB,

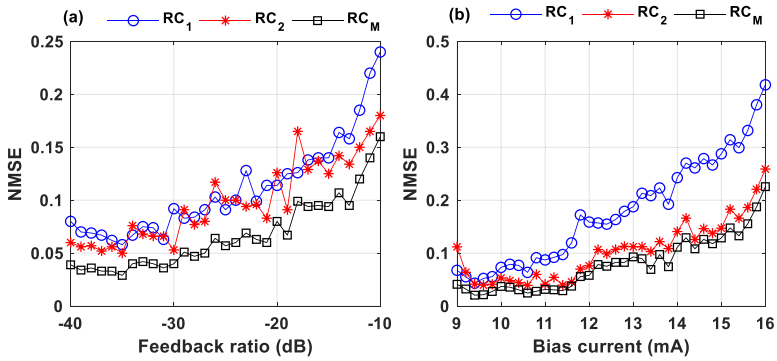


Fig. 4. (a) NMSEs as a function of feedback ratio under $j_1 = j_2 = 10$ mA, (b) NMSEs as a function of bias current under $k_1 = k_2 = -35$ dB.

and RC_M achieves the minimum value of 0.03 at $k_1 = k_2 = -35$ dB. According to the results shown in Fig. 4(a), next, we set $k_1 = k_2 = -35$ dB and investigate the variation of NMSEs with the bias current (j_1, j_2) of two SLs. As shown in Fig. 4(b), when j_1 and j_2 increase from 9 mA to 16 mA, the three curves show a similar trend, *i.e.*, the NMSEs first decrease and then gradually increase. The minimum values for RC_1 (NMSE = 0.042) and RC_2 (NMSE = 0.04) are located at $j_1 = 9.4$ mA and $j_2 = 9.6$ mA, respectively. We can find that, in one hand, under a small bias current near 9 mA, the NMSEs are relatively high, which may be caused by the relatively large spontaneous emission noises of two SLs. In the other hand, when the bias current is larger than 12 mA, the NMSEs gradually increase with the increase of bias current. The reason for this result is as follow: in this work, the V_{pp} of AWG output is fixed at 500 mv, which means that the proportion of the signal in the current is decreased with the increase of bias current. For RC_1, the NMSE is less than 0.1 in the range of 9 mA $\leq j_1 \leq$ 11.6 mA. For RC_2, the NMSE is less than 0.1 in the range of 9.2 mA $\leq j_2 \leq$12 mA. Within the tested range, the NMSE for RC_M is always lower than that for RC_1 and RC_2, and it is less than 0.1 when j_1 and j_2 are located in the wider range between 9 mA and 13 mA.

From above tests we can find that based on the merged state matrix X_M, RC_M performs better than RC_1 and RC_2, and can achieve a good performance in a wider parameter range. This is mainly due to the fact that the number of virtual node states for RC_M is doubled with respect to RC_1 and RC_2, which is beneficial for increasing the dimensionality of reservoir state space. Obviously, under such a situation, the DPR of RC_M is about 32.5 MSa/s ($=(n\theta)^{-1}$), which is same as that of RC_1 and RC_2. Next, we further investigate whether the system has a potential ability to improve DPR while maintaining a good performance. To accomplish this test without tailoring the fiber of two optical feedback loops, we adopt a method introduced in [38 that only n_r ($n_r \leq n$) points of the input stream is filled by the masked information within one period (as shown in Fig. 5(a)). Correspondingly, in the readout layer, only the virtual node states containing masked information are preserved for training and testing, and the others are discarded (as shown in Fig. 5(b)). Therefore, by changing the number n_r of virtual node states actually selected for readout, the potential DPR of the system can be evaluated, which is equal to $(n_r\theta)^{-1}$.

Fig. 6 shows the variation of NMSEs with the number n_r of virtual node states actually selected for readout, where the bias current of two SLs are set at 9.4 mA and the feedback ratio of two SLs are set at -35 dB. In this test, we increase n_r from 20 to 614, the potential DPR of the system is correspondingly decreased from 1 GSa/s to 32.5 MSa/s. For convenience, we also mark DPR on the abscissa of Fig. 6. We can find that, under a high DPR of 1 GSa/s, RC_1, RC_2, and RC_M show very poor performances since a too small n_r results in a low-dimensional state space. With the increasing of n_r, NMSEs first decrease rapidly and then tend to stable due to the saturation of state information. When n_r is larger than 300, RC_1 can achieve a NMSE stabilized below 0.1. As a consequence, the maximum DPR for RC_1 to achieve a good prediction performance is about 67 MSa/s. For RC_2, the NMSE is smaller than 0.1 when $n_r \geq 220$, which corresponds to a maximum

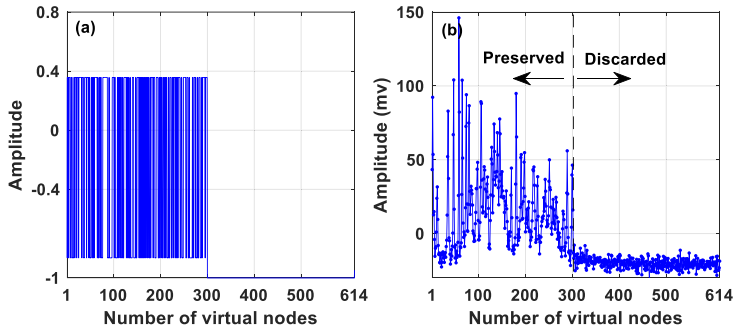


Fig. 5. Waveform of (a) masked information in one period and (b) virtual node states, where the number n_{r} of virtual node states selected for readout is 300.

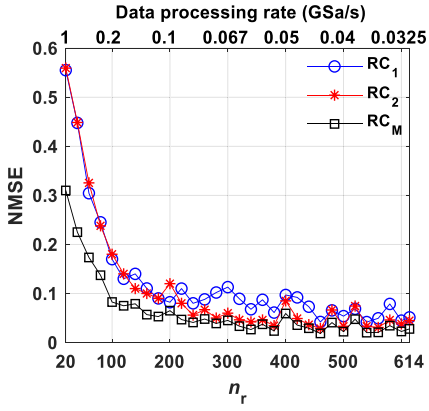


Fig. 6. NMSEs as a function of n_{r} under $j_1 = j_2 = 9.4$ mA and $k_1 = k_2 = -35$ dB.

DPR of 91 MSa/s to maintain a good prediction performance. During the variation of n_{r}, the NMSE for $\mathrm{RC_M}$ is always lower than that for $\mathrm{RC_1}$ and $\mathrm{RC_2}$, and it is lower than 0.1 when $n_{\mathrm{r}} \geq 100$. As a result, under the premise of ensuring a NMSE lower than 0.1, $\mathrm{RC_M}$ can achieve a maximum DPR of 200 MSa/s, which is about three times and twice that of $\mathrm{RC_1}$ and $\mathrm{RC_2}$, respectively. Therefore, we can find that compared with a system based on a single reservoir, the RC system based on two parallel reservoirs possesses the ability to double DPR while maintaining a good prediction performance.

3.2 Multi-Waveform Recognition Task

In this section, we use a multi-waveform recognition task to evaluate the performance of this RC system for classifying different signals. The task used in this work is similar to that used in [10] but containing more waveforms. The target of this task is to correctly classify randomly concatenated waveforms including sine, square, triangle, and such waveforms with doubled frequency (as shown in Fig. 7(a)). In the input layer, a waveform is first discretized into 25 points per period and multiplied by a mask matrix with dimension of 25 $\times$ n (n is the number of virtual nodes). Next, the masked information (as shown in Fig. 7(b)) with dimension of 1 $\times$ n is serially input into the reservoir in a period T. In the output layer, six linear classifiers are configured, and each classifier corresponds to a kind of waveform. The target function is 1 if the signal corresponds to the right classifier, and -1 if it does not (as shown in Fig. 7(c)). In the training procedure, based on the target function and virtual node states collected from two SLs, the output weight $\mathrm{W_{out}}$ can be optimized by using (1). In the testing procedure, for each input waveform, a calculated result corresponds to six classifiers

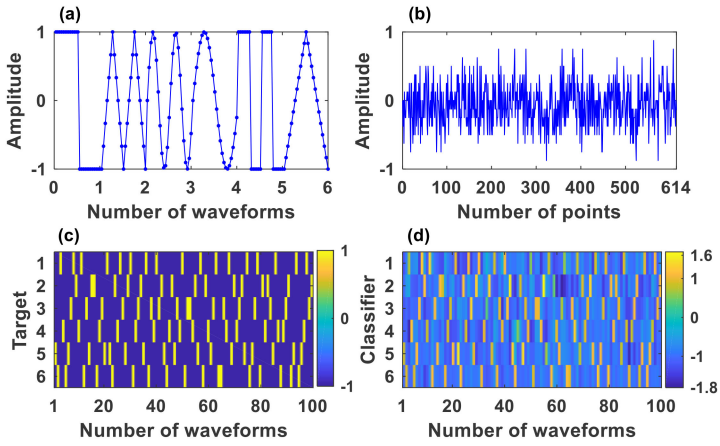


Fig. 7. Examples of multi-waveform recognition task. (a) Original input signal. (b) Masked information in one period. (c) Target function. (d) Calculated result of classifiers.

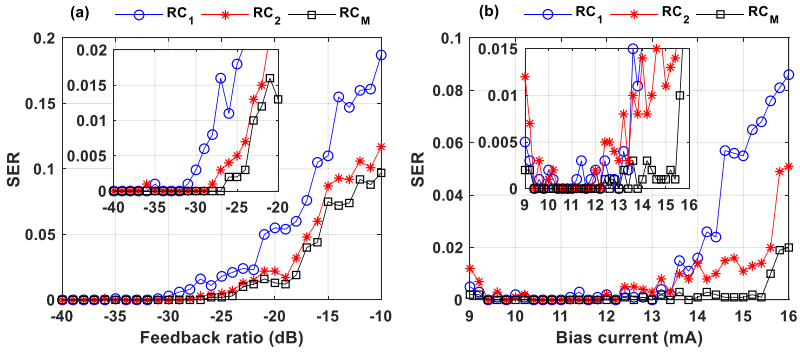


Fig. 8. (a) SERs as a function of feedback ratio under $j_1 = j_2 = 10$ mA. (b) SERs as a function of bias current under $k_1 = k_2 = -40$ dB. The inset is a part of enlarged view.

(as shown in Fig. 7(d)) can be obtained. Then, a winner-takes-all approach is applied, *i.e.*, the classifier with the maximum value is selected as the recognition result [12], [18], [37]. In this test, we use 3000 waveforms for training and 1000 waveforms for testing. The recognition performance is quantitatively evaluated by the signal error rate (SER). Referring to the decision threshold around 10^{-3} used in a similar classification task in [36], we set a moderate value of 5×10^{-3} as a reference line to evaluate the classification performance of RC_1, RC_2, and RC_M for this task. Note that the multi-waveform recognition task may be treated as a regression problem in other works, and the performance can also be evaluated by NMSE [38]. For the reference line, the value of SER is 5×10^{-3}, and the corresponding value of NMSE is about 0.23.

We first evaluate that under a DPR of 32.5 MSa/s ($n_\mathrm{r} = n = 614$), the variation of SERs for RC_1, RC_2, and RC_M under different parameters (feedback ratio and bias current) of two SLs, and the results are plotted in Fig. 8. Fig. 8(a) shows SERs as a function of feedback ratio (k_1, k_2) when two SLs are biased at 10 mA. We can see that under a feedback ratio smaller than -31 dB, the SERs of RC_1, RC_2, and RC_M are all lower than 0.005. It is worth pointing out that for this task, relatively low SERs can be achieved even under a small feedback ratio of -40 dB, which may be due to the low memory requirement of this task. As the feedback ratio increases from -30 dB to -10 dB, the SERs gradually increase, which is caused by oscillated system states under a relatively large feedback ratio. RC_1 and RC_2 can achieve SERs less than 0.005 in the range of $k_1 \leq -30$ dB and $k_2 \leq -26$ dB, respectively. RC_M can achieve a SER less than 0.005 when the feedback ratio of two

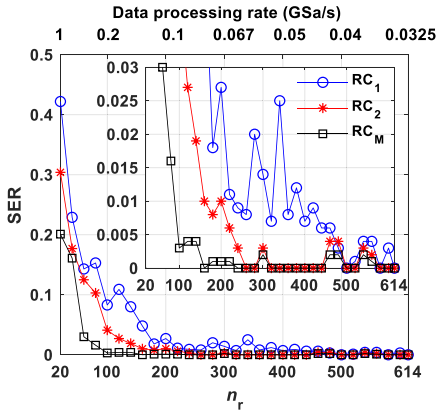


Fig. 9. SERs as a function of n_r under $j_1 = j_2 = 10$ mA and $k_1 = k_2 = -40$ dB , where the inset is a part of enlarged view.

SLs are less than -24 dB. Obviously, the range of feedback ratio for RC_M with good performances is wider than those for RC_1 and RC_2. According to the results shown in Fig. 8(a), we next set $k_1 = k_2 = -40$ dB and show the SERs as a function of bias current of two SLs in Fig. 8(b). As shown in Fig. 8(b), when two SLs are biased at low current close to their threshold current, the SERs are relatively high due to larger spontaneous emission noise of two SLs. When the two SLs are biased at current larger than 13 mA, the SERs increase significantly due to the reduced modulation response [39]. From the inset of Fig. 8(b), we can find that RC_1 and RC_2 can achieve SERs lower than 0.005 in the range of 9.2 mA $\leq j_1 \leq$ 13.4 mA and 9.4 mA $\leq j_2 \leq$ 12 mA, respectively. For RC_M, SER is lower than 0.005 when the bias current of two SLs are in a relatively wide range between 9 mA and 15.4 mA. Based on Fig. 4 and Fig. 8, one can find that RC_1, RC_2, and RC_M show similar dependence on the parameters of feedback ratio and bias current, which indicates the best RC performance prefers some particular dynamical states of the system. As demonstrated in previous studies [15], [39], for attaining good performance, RC systems are required to possess consistency and separation properties, which can be satisfied when the RC systems operate at an asymptotically stable state without input signal.

From the above test we can find that under a DPR of 32.5 MSa/s, the RC system based on two parallel reservoirs shows better classification performance and stronger parameter robustness than the system with a single reservoir. At last, we investigate whether the system can achieve a good classification performance using relatively few virtual node states, thereby possessing the ability to increase DPR. Similar to the test shown in Fig. 6, we increase the number n_r of virtual node states selected for readout from 20 to 614, so the potential DPR is correspondingly decreased from 1 GSa/s to 32.5 MSa/s. Fig. 9 shows the variation of SERs with the number n_r, where the bias current and the feedback ratio of two SLs are fixed at 10 mA and -40 dB, respectively. It can be seen that as n_r increases from 20 to 614, the SERs for RC_1, RC_2 and RC_M drop sharply first and then tend to saturation. For RC_1 and RC_2, SERs drop to below 0.005 when n_r is larger than 480 and 240, respectively. It indicates that the maximum DPR for RC_1 and RC_2 to realize a good classification performance (SER $\leq$ 0.005) are 41.7 MSa/s and 83.3 MSa/s, respectively. For RC_M, the SER decreases to below 0.005 when n_r is larger than 100, which corresponds to a maximum DPR of 200 MSa/s for maintaining a good performance. As a result, for this multi-waveform recognition task, it is feasible to double the DPR by using the RC system based on two parallel reservoirs.

In addition, we have inspected the performance of RC_M under other integration schemes for constructing $\boldsymbol{x}^i_M(n)$ such as $\boldsymbol{x}^i_M(n) = [x^i_2(1), x^i_2(2), \ldots, x^i_2(n_r), x^i_1(1), x^i_1(2), \ldots, x^i_1(n_r)]^T$ or $\boldsymbol{x}^i_M(n) = [x^i_1(1), x^i_2(1), x^i_1(2), x^i_2(2), \ldots, x^i_1(n_r), x^i_2(n_r)]^T$, and the results demonstrate that RC_M always shows better performance than RC_1 and RC_2.

4. Conclusion

In this work, we experimentally investigate the performance of an optical RC system based on two parallel time-delay reservoirs which are constituted of two distributed feedback SLs (SL_1, SL_2) subject to optical feedback, respectively. The masked information generated by an AWG is first split into two parts and then injected into two reservoirs through directly modulating the pump current of two SLs. The output intensities of SL_1 and SL_2 are sampled as virtual node states and form two state matrixes X_1 and X_2, respectively. As an RC based on two parallel reservoirs, we perform RC (named as RC_M) using the state matrixes X_M (merged by X_1 and X_2) for training and testing. Meanwhile, the performance of RC_M is compared with that of RC_1 and RC_2, where RC_1 and RC_2 are performed using X_1 and X_2 for training and testing, respectively. By employing Santa Fe time series prediction task and multi-waveform recognition task, the performances of RC_1, RC_2 and RC_M are investigated under different parameters of two SLs. The results show that under a DPR of 32.5 MSa/s, for both tasks, RC_M shows better performance and stronger parameter robustness than RC_1 and RC_2. Furthermore, through analyzing the dependence of the system performance on the number n_r of the selected virtual node states in readout layer, the potential DPR of the system is investigated. For Santa Fe time series prediction task, under the premise of ensuring NMSE less than 0.1, the potential DPR for RC_1, RC_2, and RC_M can achieve 67 MSa/s, 91 MSa/s, and 200 MSa/s, respectively. For multi-waveform recognition task, under the premise of ensuring SER lower than 0.005, the potential DPR for RC_1, RC_2, and RC_M can achieve 40 MSa/s, 90 MSa/s, and 200 MSa/s, respectively. It is evident that for both tasks, the proposed RC system can double the DPR while maintaining a good performance compared with the system based on a single time-delay reservoir. Although the maximum potential DPR (200 MSa/s) achieved in this work does not reach to the state-of-the-art level (about 800 MSa/s) [22], [23], the results demonstrate the feasibility of using an RC system with two parallel time-delay reservoirs to enable fast information processing.

References

[1] F. Musumeci *et al.*, "An overview on application of machine learning techniques in optical networks," *IEEE Commun. Surv. Tut.*, vol. 21, no. 2, pp. 1383–1408, Nov. 2019.

[2] H. Jaeger and H. Haas, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, no. 2, pp. 78–80, Apr. 2004.

[3] W. Maass, T. Natschläger, and H. Markram, "Real-time computing without stable states: A new framework for neural computation based on perturbations," *Neural Comput.*, vol. 14, no. 11, pp. 2531–2560, Nov. 2002.

[4] D. Verstraeten, B. Schrauwen, M. D'Haene, and D. Stroobandt, "An experimental unification of reservoir computing methods," *Neural Netw.*, vol. 20, no. 3, pp. 391–403, Apr. 2007.

[5] M. C. Soriano, D. Brunner, M. Escalona-Morán, C. R. Mirasso, and I. Fischer, "Minimal approach to neuro-inspired information processing," *Front. Comput. Neurosci.*, vol. 9, Jun. 2015, Art. no. 68.

[6] K. Vandoorne, J. Dambre, D. Verstraeten, B. Schrauwen, and P. Bienstman, "Parallel reservoir computing using optical amplifiers," *IEEE Trans. Neural Netw.*, vol. 22, no. 9, pp. 1469–1481, Sep. 2011.

[7] L. Appeltant *et al.*, "Information processing using a single dynamical node as complex system," *Nat. Commun.*, vol. 2, Sep. 2011, Art. no. 468.

[8] M. C. Soriano *et al.*, "Delay-based reservoir computing: Noise effects in a combined analog and digital implementation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 2, pp. 388–393, Feb. 2015.

[9] L. Larger *et al.*, "Photonic information processing beyond turing: An optoelectronic implementation of reservoir computing," *Opt. Exp.*, vol. 20, no. 3, pp. 3241–3249, Jan. 2012.

[10] Y. Paquot *et al.*, "Optoelectronic reservoir computing," *Sci. Rep.*, vol. 2, no. 1, Feb. 2012, Art. no. 287.

[11] M. C. Soriano *et al.*, "Optoelectronic reservoir computing: Tackling noise-induced performance degradation," *Opt. Exp.*, vol. 21, no. 1, pp. 12–20, Jan. 2013.

[12] Q. C. Zhao, H. X. Yin, and H. G. Zhu, "Simultaneous recognition of two channels of optical packet headers utilizing reservoir computing subject to mutual-coupling optoelectronic feedback," *Optik*, vol. 157, pp. 951–956, 2018.

[13] Y. P. Chen *et al.*, "Reservoir computing system with double optoelectronic feedback loops," *Opt. Exp.*, vol. 27, no. 20, pp. 27431–27440, Sep. 2019.

[14] F. Duport, B. Schneider, A. Smerieri, M. Haelterman, and S. Massar, "All-optical reservoir computing," *Opt. Exp.*, vol. 20, no. 20, pp. 22783–22795, Sep. 2012.

[15] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, "Parallel photonic information processing at gigabyte per second data rates using transient states," *Nat. Commun.*, vol. 4, Jan. 2013, Art. no. 1364.

[16] J. Vatin, D. Rontani, and M. Sciamanna, "Experimental realization of dual task processing with a photonic reservoir computer," *APL Photon.*, vol. 5, no. 8, Aug. 2020, Art. no. 086105.

[17] Q. Vinckier *et al.*, "High performance photonic reservoir computer based on a coherently driven passive cavity," *Optica*, vol. 2, no. 5, pp. 438–446, May 2015.
[18] X. R. Bao, Q. C. Zhao, H. X. Yin, and J. Qin, "Recognition of the optical packet header for two channels utilizing the parallel reservoir computing based on a semiconductor ring laser," *Mod. Phys. Lett. B*, vol. 32, no. 14, May 2018, Art. no. 1850150.
[19] Q. Cai *et al.*, "Modulation format identification in fiber communications using single dynamical node-based photonic reservoir computing," *Photon. Res.*, vol. 9, no. 1, pp. B1–B8, Jan. 2021.
[20] R. M. Nguimdo, G. Verschaffelt, J. Danckaert, and S. G. Van, "Fast photonic information processing using semiconductor lasers with delayed optical feedback: Role of phase dynamics," *Opt. Exp.*, vol. 22, no. 7, pp. 8762–8686, Apr. 2014.
[21] Y. Kuriki, J. Nakayama, K. Takano, and A. Uchida, "Impact of input mask signals on delay-based photonic reservoir computing with semiconductor lasers," *Opt. Exp.*, vol. 26, no. 5, pp. 5777–5788, Mar. 2018.
[22] K. Takano *et al.*, "Compact reservoir computing with a photonic integrated circuit," *Opt. Exp.*, vol. 26, no. 22, pp. 29424–29439, Oct. 2018.
[23] K. Harkhoe, G. Verschaffelt, A. Katumba, P. Bienstman, and G. V. D. Sande, "Demonstrating delay-based reservoir computing using a compact photonic integrated chip," *Opt. Exp.*, vol. 28, no. 3, pp. 3086–3096, Feb. 2020.
[24] S. Ortín and L. Pesquera, "Reservoir computing with an ensemble of time-delay reservoirs," *Cogn. Comput.*, vol. 9, no. 3, pp. 327–336, Jun. 2017.
[25] A. Rohm and K. Ludge, "Multiplexed networks: Reservoir computing with virtual and real nodes," *J. Phys. Commun.*, vol. 2, no. 8, Aug. 2018, Art. no. 085007.
[26] X. X. Guo, S. Y. Xiang, Y. H. Zhang, L. Lin, A. J. Wen, and Y. Hao, "Four-channels reservoir computing based on polarization dynamics in mutually coupled VCSELs system," *Opt. Exp.*, vol. 27, no. 16, pp. 23293–23306, Aug. 2019.
[27] C. Sugano, K. Kanno, and A. Uchida, "Reservoir computing using multiple lasers with feedback on a photonic integrated circuit," *IEEE J. Sel. Top. Quantum Electron.*, vol. 26, no. 1, Jan. 2020, Art. no. 1500409.
[28] X. R. Bao, Q. C. Zhao, and H. X. Yin, "A multiple-input multiple-output reservoir computing system subject to optoelectronic feedbacks and mutual coupling," *Entropy*, vol. 22, no. 2, Feb. 2020, Art. no. 231.
[29] A. Bogris, C. Mesaritakis, S. Deligiannidis, and P. Li, "Fabry-Perot lasers as enablers for parallel reservoir computing," *IEEE J. Sel. Top. Quantum Electron.*, vol. 27, no. 2, Mar. 2021, Art. no. 7500307.
[30] Y. S. Hou, G. Q. Xia, E. Jayaprasath, D. Z. Yue, W. Y. Yang, and Z. M. Wu, "Prediction and classification performance of reservoir computing system using mutually delay-coupled semiconductor lasers," *Opt. Commun.*, vol. 433, pp. 215–220, Feb. 2019.
[31] D. Z. Yue, Z. M. Wu, Y. S. Hou, C. X. Hu, Z. F. Jiang, and G. Q. Xia, "Reservoir computing based on two parallel reservoirs under identical electrical message injection," *IEEE Photon. J.*, vol. 13, no. 1, Feb. 2021, Art. no. 7800311.
[32] X. X. Guo, S. Y. Xiang, Y. H. Zhang, L. Lin, A. J. Wen, and Y. Hao, "High-speed neuromorphic reservoir computing based on a semiconductor nanolaser with optical feedback under electrical modulation," *IEEE J. Sel. Top. Quantum Electron.*, vol. 26, no. 5, Sep. 2020, Art. no. 1500707.
[33] N. Q. Li *et al.*, "Multiscale ordinal symbolic analysis of the Lang-Kobayashi model for external-cavity semiconductor lasers: A test of theory," *IEEE J. Quantum Electron.*, vol. 51, no. 8, Aug. 2015, Art. no. 2200206.
[34] U. Hübner, N. B. Abraham, and C. O. Weiss, "Dimensions and entropies of chaotic intensity pulsations in a single-mode far-infrared NH3 laser," *Phys. Rev. A*, vol. 40, no. 11, pp. 6354–6365, Dec. 1989.
[35] D. Z. Yue *et al.*, "Performance optimization research of reservoir computing system based on an optical feedback semiconductor laser under electrical information injection," *Opt. Exp.*, vol. 27, no. 14, pp. 19931–19939, Jul. 2019.
[36] A. Argyris, J. Bueno, and I. Fisher, "Photonic machine learning implementation for signal recovery in optical communications," *Sci. Rep.*, vol. 8, May 2018, Art. no. 8487.
[37] L. Larger, A. Baylon-Fuentes, R. Martinenghi, V. S. Udaltsov, Y. K. Chembo, and M. Jacquot, "High-speed photonic reservoir computing using a time-delay-based architecture: Million words per second classification," *Phys. Rev. X*, vol. 7, no. 1, Feb. 2017, Art. no. 011015.
[38] X. X. Guo, S. Y. Xiang, Y. H. Zhang, L. Lin, A. J. Wen, and Y. Hao, "Polarization multiplexing reservoir computing based on a VCSEL with polarized optical feedback," *IEEE J. Sel. Top. Quantum Electron.*, vol. 26, no. 1, Jan./Feb. 2020, Art. no. 1700109.
[39] D. Z. Yue, Z. M. Wu, Y. S. Hou, and G. Q. Xia, "Effects of some operation parameters on the performance of a reservoir computing system based on a delay feedback semiconductor laser with information injection by current modulation," *IEEE Access*, vol. 7, pp. 128767–128773, Aug. 2019.