Optics Letters

Multi-task photonic time-delay reservoir computing based on polarization modulation

LONG HUANG AND JIANPING YAO*

Microwave Photonics Research Laboratory, School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada
**Corresponding author: jpyao@eecs.uottawa.ca*

Received 26 October 2022; revised 17 November 2022; accepted 17 November 2022; posted 22 November 2022; published 13 December 2022

We propose and experimentally demonstrate a multi-task photonic time-delay reservoir computing (RC) system based on polarization modulation. The key component in the system is a polarization modulator (PolM) that functions, jointly with a polarization controller (PC) and a polarizer, as an equivalent Mach–Zehnder modulator (MZM) to perform electrical to optical conversion and to provide nonlinear operation. By adjusting the bias of the equivalent MZM, the nonlinear function can be optimized for different tasks to achieve the best multi-task performance. In this paper, the task-independent information processing capacity (IPC) of the time-delay RC system is evaluated. The results show that the readout bias of the equivalent MZM leads to a different IPC which can be optimized for different tasks. Two benchmark tasks (NARMA10 and IPIX radar signal prediction) are performed experimentally. The readout bias is adjusted independently for each of the two tasks to give a minimum normalized mean square error (NMSE), which are 0.2103 and 0.0031 for the NARMA10 and IPIX radar signal prediction tasks at a speed of 1.06 Mb/s, respectively.

https://doi.org/10.1364/OL.479472

The term reservoir computing (RC) was coined by Verstraeten *et al.* [1] in 2007, which was proposed to unify two closely related recurrent neural network (RNN) structures, the echo state network by Jaeger [2] proposed in 2001, and the liquid state machine by Maass *et al.* [3] proposed in 2002. In an RC system, input signals are fed into a fixed RNN called a reservoir. A simple readout layer is trained to read the state of the reservoir and map it to a desired output. The benefit of this framework is that the training is performed only at the readout stage and the laborious process of gradient-descent RNN training is avoided. This simplification makes it possible to implement RC systems based on photonics to provide a practical yet more powerful hardware platform for RC. A photonic RC system can be implemented based on coupled semiconductor optical amplifiers (OSAs) [4], an integrated passive silicon photonic circuit with the nonlinearity achieved at the readout layer [5], and a semiconductor laser network based on diffractive coupling [6]. All these implementations rely on a real network with photonic nodes that are spatially distributed. However, the concept of a virtual network for RC using a time-delay feedback system in electronics has been proposed [7]. In a delay-based RC system, input signals are time-multiplexed and injected into the reservoir via a nonlinear node. Then, the transient temporal dynamics of the reservoir system are sampled and considered as virtual network nodes. The time-delay feedback concept drastically simplifies the hardware implementation and was later extended to photonic implementations. Photonic time-delay RC systems have been demonstrated using optical modulators [8], semiconductor optical amplifiers [9], and semiconductor lasers [10]. To process multiple tasks simultaneously, multiplexing of the two polarization modes in vertical-cavity surface-emitting lasers (VCSELs) were explored for parallel computing [11]. In [12], two directional modes in a single ring laser are used to process two tasks simultaneously. The approaches reported in [11] and [12] are all-optical RC schemes, where optical-to-electrical conversion and nonlinear functions are implemented separately. Parallel RC can also be implemented in optoelectronic systems where optical-to-electrical conversion and the nonlinear function are achieved in a single device [13]. Since optimal nonlinear functions are different for different tasks, the shared nonlinear node cannot achieve the best performance for different tasks simultaneously [13].

In this paper, a photonic RC system for multi-task computing based on polarization modulation is proposed and experimentally demonstrated. First, a polarization modulator (PolM) is employed to convert the feedback electrical signal to a polarization-modulated optical signal. The PolM is operating, jointly with a polarization controller (PC) and a polarizer, to function as an equivalent Mach–Zehnder modulator (MZM). The polarization-modulated signal is converted to an intensity-modulated optical signal at the output of the polarizer. By tuning the PC, the readout bias of the equivalent MZM is changed, which would lead to different nonlinear functions. Due to the structure of the equivalent MZM, the nonlinear function can be optimized for different tasks to simultaneously achieve the best performance. The information processing capacity (IPC) of the time-delay RC performance on the bias is evaluated numerically. It shows that the tuning of the readout bias of the equivalent MZM would lead to a different IPC which can be adapted to different tasks to achieve the best multi-task performance. Then, two benchmark tasks, NARMA10 and IPIX radar signal prediction, are performed experimentally. The bias of the equivalent

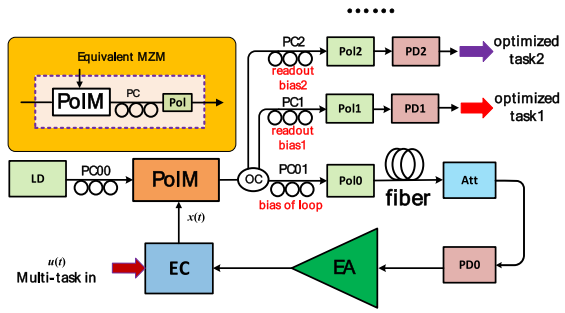

Fig. 1. Schematic diagram of the proposed photonic RC system.

MZM is adjusted separately for the two tasks to give the minimum normalized mean square errors (NMSEs) of 0.2103 and 0.0031 at a speed of 0.53 Mb/s for a single task and 1.06 Mb/s for two tasks.

The proposed photonic RC system is shown in Fig. 1. The key component in the system is the PolM, which is a special phase modulator that supports phase modulation for both the transverse-electric (TE) and transverse-magnetic (TM) modes with opposite phase modulation indices [14]. When an electrical signal $s(t)$ is applied to the PolM, the phase-modulated signals after the PolM can be written as

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix} = \begin{bmatrix} e^{j\gamma s(t)} \\ e^{-j\gamma s(t)} \end{bmatrix} e^{j\omega_c t}, \tag{1}$$

where γ is the phase modulation index and ω_c is the center frequency of the optical carrier. The PolM operates jointly with a PC and a polarizer to form an equivalent MZM. The structure of a PolM-based equivalent MZM is shown in the inset of Fig. 1. By adjusting the PC, the optical signal at the output of the polarizer is given by

$$E_o(t) = \left(\frac{\sqrt{2}}{2}\right)(E_x e^{j2\varphi} + E_y), \tag{2}$$

where 2φ is a phase shift introduced by the PC between E_x and E_y. Then, the optical intensity can be written as

$$I_o(t) = |E_o|^2 \propto \cos^2[\gamma s(t) + \varphi]. \tag{3}$$

As can be seen from Eq. (3), the polarization-modulated signal is converted to an intensity-modulated signal at the output of the polarizer, and the phase of the signal can be adjusted by adjusting the bias through tuning the PC.

As can be seen from Fig. 1, a light wave generated by a laser diode (LD) is sent to a PolM via a PC (PC00). The light wave after the PolM is split into multiple branches by a $1 \times N$ optical coupler (OC). One branch is for feedback to close the optoelectronic loop. Each of the other branches is for readout of a specific task. In the feedback loop, the PolM operates jointly with a PC (PC01) and a polarizer (Pol0) to form an equivalent MZM. Then, the optical signal at the output of the equivalent MZM is delayed by a spool of fiber, attenuated by an attenuator (Att), and sent to a photodetector (PD0) for optical-to-electrical conversion. The generated electrical signal is amplified by an electrical amplifier (EA), combined with the multi-task input signal at an electrical coupler (EC), and applied to the PolM via the RF port.

The dynamics of the proposed RC system is analyzed. We start with a dimensionless variable $x(t) = \pi V(t)/2V_\pi$, where $V(t)$ is a voltage signal driving the equivalent MZM with a half-voltage of V_π. The feedback circuitry is assumed to be a second-order bandpass filter formed by a first-order low-pass filter with a 3-dB cutoff frequency f_L and a first-order high-pass filter with a 3-dB cutoff frequency f_H. Like a conventional MZM, the equivalent MZM has a nonlinear transfer function $\cos^2(.)$. A differential equation characterizing the state $x(t)$ of RC system is given by

$$x(t) + \tau\frac{dx(t)}{dt} + \frac{1}{\sigma}\int_{t_0}^{t} x(s)ds = \beta\cos^2[x(t-T) + \alpha u(t) + \psi_0], \tag{4}$$

where $\tau = 1/(2\pi f_H)$ and $\sigma = 1/(2\pi f_L)$ are the characteristic response times associated with the high and low cutoff frequencies f_H and f_L, respectively, β denotes the feedback strength, $u(t)$ is the input signal to the RC system, α is a scaling factor, ψ_0 is the bias of the equivalent MZM, and T stands for the feedback time delay. The signal at the output of the RC system is given by

$$y(t) = \cos^2[x(t) + \psi], \tag{5}$$

where ψ is the bias of the equivalent MZM at a readout layer. The Taylor expansion of $y(t)$ at the bias ψ is given by

$$\begin{aligned} y(t) = {} & \frac{1}{2} + \frac{1}{2}\cos(2\psi)\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}[2x(t)]^{2n} \\ & - \frac{1}{2}\sin(2\psi)\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}[2x(t)]^{2n+1} \end{aligned}. \tag{6}$$

Therefore, we can adjust ψ to achieve different nonlinearity in a readout layer, eventually leading to a different IPC.

Next, we perform a numerical simulation to show the dependence of the IPC on the bias ψ. The IPC introduced in [15] is widely used as a task-independent measure to characterize the processing performance of an RC system. In the evaluation of IPC, the input signal is a random sequence taking arbitrary values from -1 to 1, and the corresponding state of the RC system is recorded. The target signal is given by $z(t) = \prod_i P_{d_i}(u(t-i))$, where $P_{di}(.)$ is the normalized Legendre polynomial of degree d_i at delay i. The IPC of an RC system to reconstruct the function z from its state is quantified in terms of the mean square error (MSE) $MSE[\hat{z}] = \sum_{t=1}^{T}(\hat{z}(t) - z(t))^2/T$, where $\hat{z}(t) = \sum_{i=1}^{N} W_i x_i(t) + W_0$ is a linear estimator of $z(t)$ by the state variables $x_i(t)$. Then, the IPC of an RC system to reconstruct the function z is defined as $C = 1 - \min MSE[\hat{z}]/\langle z^2\rangle$, where $\langle z^2\rangle = \sum_{t=1}^{T} z(t)^2/T$. The minimum MSE is calculated by the linear regression to obtain the IPC. The IPC of the proposed RC system as a function of the readout bias is numerically calculated by the fourth-order Runge–Kutta method. In the simulation, the interval of the virtual nodes θ is set to $\theta = 50$ ps, the time delay T is set to $T = 20\theta$, the bias of the equivalent MZM in the loop ψ_0 is set to $\psi_0 = \pi/4$, the characteristic response times associated with the high and low cutoff frequencies τ and σ are set to $\tau = 19.89$ ps and $\sigma = 51.34$ ps, respectively. The calculated IPC versus the readout bias ψ is given in Fig. 2. As can be seen, when $\psi = 0$, the system only has an IPC of even degrees. This phenomenon can be explained from the Taylor expansion of $y(t)$. When $\psi = 0$, the Taylor series of $y(t)$ is given by $y(t) = 1 - x^2(t) + (1/3)x^4(t) + (2/45)x^6 + \ldots$. As can be seen, $y(t)$ only consists of terms of even powers of $x(t)$. We can also see from Fig. 2 that when $\psi = \pi/4$, the RC only has an IPC of odd degrees. This phenomenon can also be explained from the

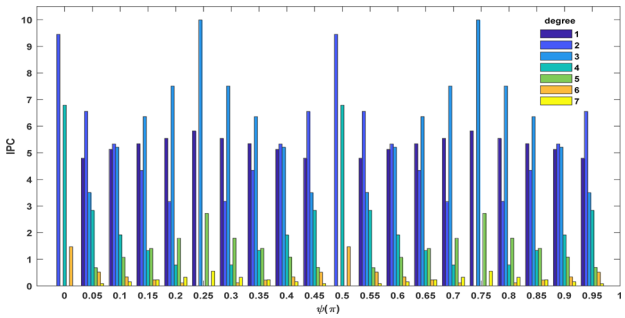

Fig. 2. IPC evaluated under different readout bias until a degree of 7. The capacities for different degrees are plotted by different colors.

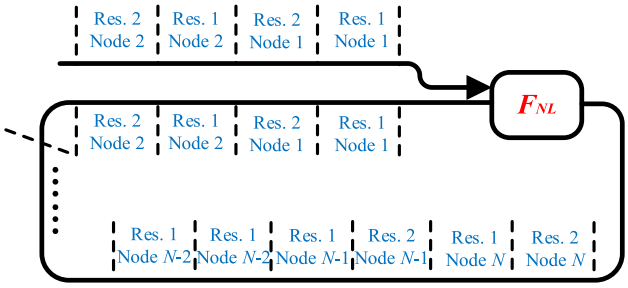

Fig. 3. Schematic to show the time-multiplexing of the two tasks.

Taylor series of $y(t)$ for $\psi = \pi/4$, $y(t) = 1/2 - x(t) + (2/3)x^3(t) - (2/15)x^5(t) + (4/315)x^7(t) + \ldots$. As can be seen, $y(t)$ only consists of terms of odd powers of $x(t)$. Based on the dependence of the IPC on the readout bias ψ, the readout bias can be adjusted to adapt to different tasks with different nonlinearity. The IPC is periodic with a period of 0.5π since the Taylor expansion of $y(t)$ at the bias $\psi + 0.5\pi$ is given by

$$y(t) = \frac{1}{2} - \frac{1}{2}\cos(2\psi)\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}[2x(t)]^{2n} + \frac{1}{2}\sin(2\psi)\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}[2x(t)]^{2n+1}. \quad (7)$$

Compare Eq. (7) with Eq. (6), the absolute value of the coefficients of the Taylor series is periodic with a period of 0.5π, leading to the same period of the IPC.

The input signal is constructed by time-multiplexing of several tasks, as introduced in [13]. To time interleave these tasks, one needs to inject the masked input signal successively by virtual nodes for the reservoir computer, as is shown in Fig. 3.

A proof-of-concept experiment is performed based on the setup shown in Fig. 1. Note that to simplify the implementation, only one readout branch is constructed in the experiment. Since the bias of each readout layer is independently controllable, the task of a specific layer can be optimized independently without affecting the performance of the other layers.

In the experiment, a tunable laser source (Keysight N7714A) is used to generate a light wave at 1550 nm with a power of 9 mW. The light wave is sent to the PolM (Vesawave) with a bandwidth of 40 GHz and V_π of 5.3 V. The optical signal after the PolM is split by a 1×2 50:50 optical coupler. Each of the two branches is connected to a PC and a polarizer to achieve an equivalent MZM. The time delay of the reservoir is 2.05 μs. The EA inside

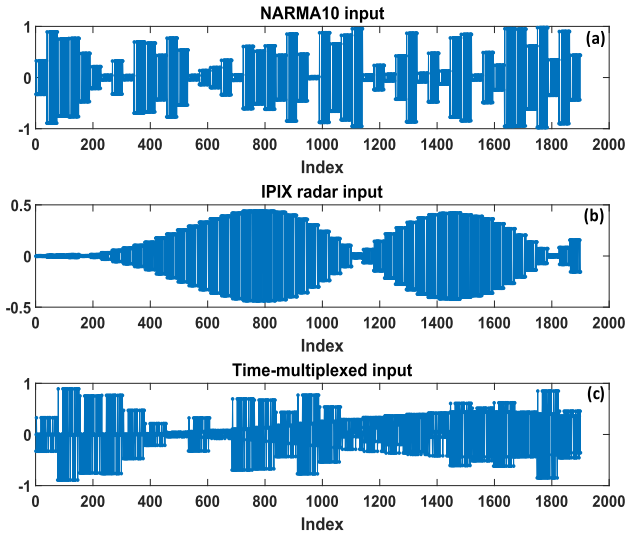

Fig. 4. Input signals to the reservoir computer. (a) NARMA10 task, (b) IPIX radar signal prediction task, and (c) time-multiplexed tasks.

the reservoir has a gain of 35 dB with a bandwidth from 30 kHz to 11 GHz. The input signal sent to the reservoir is generated by an arbitrary waveform generator (Tektronix AWG7102) with a sampling rate of 1.6 GSa/s. The signal at the output of each readout branch is also sent to a PD with a bandwidth of 10 GHz for optical to electrical conversion. The output electrical signal is sampled by an oscilloscope (Agilent DSO-X 93204A) with a sampling rate of 200 MSa/s and 12-bit resolution. The interval of the virtual nodes is set to 0.05 μs and the number of virtual nodes is set to 38.

First, two tasks, NARMA10 and IPIX radar signal prediction tasks, are generated and time multiplexed. The model to generate the NARMA10 series is given by the recurrence $y(n+1) = 0.3y(n) + 0.05y(n)\sum_{i=0}^{9} y(n-i) + 1.5u(n-9)u(n) + 0.1$, where $u(n)$ is a sequence of random inputs drawn from a uniform distribution over an interval [0, 0.5]. For the IPIX radar signal prediction task, we consider a radar signal backscattered from the ocean surface collected by the MacMaster University IPIX radar [13]. This task is to predict the radar signal one step in the future. Two random binary sequences with values drawn from $\{-1, 1\}$ are used as the mask signals for the two tasks. The input signals of the two tasks are mixed with the masks and time multiplexed. Then, the time-multiplexed signal is fed to the reservoir computer for processing. Figure 4 shows the input signals mixed with the masks and the time-multiplexed signal. For both tasks, 700 samples are used for training while 700 samples are used for testing. In both stages, the first 200 samples are discarded due to the need for warmup of the reservoir computer. In the training phase, the readout weights W are trained by a ridge regression algorithm and fixed in the test phase.

The feedback strength of the reservoir computer and the internal bias ψ_0 are adjusted to be at the edge of a chaos for good performance [16]. Then, the readout bias is adjusted to minimize the NMSE of the NARMA10 task, which is done by adjusting ψ. In this case, the NMSE of the NARMA10 task is 0.2103 on average. The NMSE of the IPIX radar signal prediction task is also calculated, which is 0.0111 on average. Figure 5

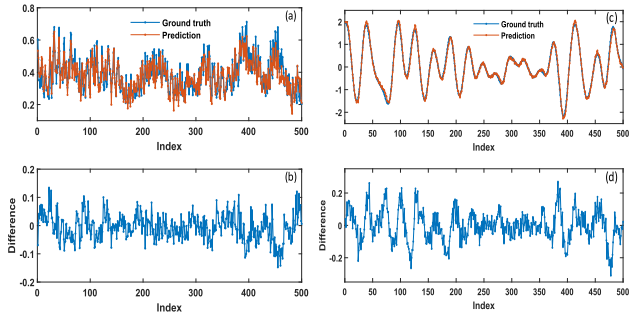

Fig. 5. Readout bias adjusted to minimize the NMSE of the NARMA10 task. (a) True and predicted NARMA10 time series and (b) the difference between them. (c) True and predicted IPIX radar signals and (d) the difference between them.

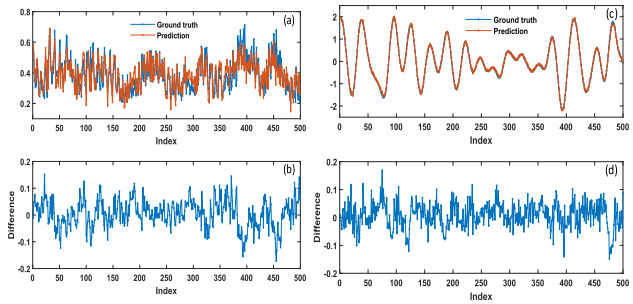

Fig. 6. Readout bias adjusted to minimize the NMSE of the IPIX radar signal prediction task. (a) True and predicted NARMA10 time series and (b) the difference between them. (c) True and predicted IPIX radar signals and (d) the difference between them.

shows the experimental results, in which Fig. 5(a) shows the true and predicted NARMA10 time series and Fig. 5(b) shows the difference between them. Figure 5(c) shows the true and predicted IPIX radar signals and Fig. 5(d) shows the difference between them. Then, ψ is adjusted to minimize the NMSE of the IPIX radar signal prediction task. The NMSE is calculated to be 0.0031 on average. The NMSE of the NARMA10 task is also calculated which is 0.2560 on average. Figure 6 shows the experimental results in which Fig. 6(a) shows the true and predicted NARMA10 time series and Fig. 6(b) shows the difference between them. Figure 6(c) shows the true and predicted IPIX radar signals and Fig. 6(d) shows the difference between them.

In conclusion, we have proposed and demonstrated a novel reservoir computer for multi-task computing based on a PolM. The key device in the system was the PolM, which could function, jointly with a PC and a polarizer, as an equivalent MZM. The performance of the reservoir computer can be optimized by adjusting hyper-parameters including the number of nodes N, the feedback strength β, the scaling factor of the input signal α, the high and low cutoff frequencies f_H and f_L, and the bias ψ_0 and ψ. Within these hyper-parameters, the readout bias ψ can be adjusted by the PolM-based equivalent MZM in each readout layer independently. Thus, optimized multi-tasks can be implemented. Numerical simulation showed by controlling the readout bias, the IPC could be adjusted. Thanks to this property, the readout bias of each readout layer can be independently adjusted to optimize the performance of multiple tasks. The proposed system to process two tasks were experimentally demonstrated. Experiment results have confirmed that the performance of each task could be independently optimized by adjusting the bias of the specific readout layer.

Funding. Natural Sciences and Engineering Research Council of Canada.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

REFERENCES

1. D. Verstraeten, B. Schrauwen, M. D'Haene, and D. Stroobandt, Neural Netw. **20**, 391 (2007).
2. H. Jaeger, "The 'echo state' approach to analysing and training recurrent neural networks," Technical Report GMD Report 148, (German National Research Center for Information Technology, 2001).
3. W. Maass, T. Natschläger, and H. Markram, Neural Comput. **14**, 2531 (2002).
4. K. Vandoorne, J. Dambre, D. Verstraeten, B. Schrauwen, and P. Bienstman, IEEE Trans. Neural Netw. **22**, 1469 (2011).
5. K. Vandoorne, P. Mechet, T. Van Vaerenbergh, M. Fiers, G. Morthier, D. Verstraeten, B. Schrauwen, J. Dambre, and P. Bienstman, Nat. Commun. **5**, 3541 (2014).
6. D. Brunner and I. Fischer, Opt. Lett. **40**, 3854 (2015).
7. 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).
8. Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, and S. Massar, Sci. Rep. **2**, 287 (2012).
9. F. Duport, B. Schneider, A. Smerieri, M. Haelterman, and S. Massar, Opt. Express **20**, 22783 (2012).
10. D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, Nat. Commun. **4**, 1364 (2013).
11. J. Vatin, D. Rontani, and M. Sciamanna, APL Photonics **5**, 086105 (2020).
12. R. M. Nguimdo, G. Verschaffelt, J. Danckaert, and G. Van der Sande, IEEE Trans. Neural Netw. Learning Syst. **26**, 3301 (2015).
13. F. Duport, A. Smerieri, A. Akrout, M. Haelterman, and S. Massar, J. Lightwave Technol. **34**, 2085 (2016).
14. J. D. Bull, N. A. F. Jaeger, H. Kato, M. Fairburn, A. Reid, and P. Ghanipour, Proc. SPIE **5577**, 133 (2004).
15. J. Dambre, D. Verstraeten, B. Schrauwen, and S. Massar, Sci. Rep. **2**, 514 (2012).
16. M. Inubushi, K. Yoshimura, Y. Ikeda, and Y. Nagasawa, *Reservoir Computing* (Springer, 2021), pp. 97–116.