Neuromorphic Computing and Engineering

PAPER • OPEN ACCESS

Hardware optimization for photonic time-delay reservoir computer dynamics

To cite this article: Meng Zhang *et al* 2023 *Neuromorph. Comput. Eng.* **3** 014008

IOP Publishing Neuromorph. Comput. Eng. 3 (2023) 014008 https://doi.org/10.1088/2634-4386/acb8d7

NEUROMORPHIC
Computing and Engineering

CrossMark

OPEN ACCESS

RECEIVED
26 October 2022

REVISED
31 January 2023

ACCEPTED FOR PUBLICATION
3 February 2023

PUBLISHED
22 February 2023

PAPER

Hardware optimization for photonic time-delay reservoir computer dynamics

Meng Zhang, Zhizhuo Liang and Z Rena Huang*

Department of Electrical Computer and System Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, United States of America
* Author to whom any correspondence should be addressed.

E-mail: huangz3@rpi.edu

Keywords: photonic time-delay reservoir computing, bifurcation dynamics, hyperparameter optimization

Supplementary material for this article is available online

Abstract

Reservoir computing (RC) is one kind of neuromorphic computing mainly applied to process sequential data such as time-dependent signals. In this paper, the bifurcation diagram of a photonic time-delay RC system is thoroughly studied, and a method of bifurcation dynamics guided hardware hyperparameter optimization is presented. The time-evolution equation expressed by the photonic hardware parameters is established while the intrinsic dynamics of the photonic RC system is quantitively studied. Bifurcation dynamics based hyperparameter optimization offers a simple yet effective approach in hardware setting optimization that aims to reduce the complexity and time in hardware adjustment. Three benchmark tasks, nonlinear channel equalization (NCE), nonlinear auto regressive moving average with 10th order time lag (NARMA10) and Santa Fe laser time-series prediction tasks are implemented on the photonic delay-line RC using bifurcation dynamics guided hardware optimization. The experimental results of these benchmark tasks achieved overall good agreement with the simulated bifurcation dynamics modeling results.

1. Introduction

Reservoir computing (RC), a subset of the recurrent neural network has attracted considerable attentions due to its simplicity in network training. A typical RC network consists of three layers: an input layer, a reservoir layer (hidden layer) and an output layer. The neuron connections in the reservoir layer are sparse and random but fixed while the training takes place only at the output layer by computing the output weight matrix. RC is a simple yet powerful machine learning approach and particularly effective processing sequential data in classification or prediction tasks such as time series forecasting [1], pattern classification [2–4], spoken digit recognition [5], signal recovery in communication systems [6, 7].

Recently, driven by edge computing applications, RC algorithms have been explored to harness on various hardware platforms such as electronic RC by nonlinear analog electronic circuits [8, 9] and FPGAs [10], reservoir by spintronics [11, 12], photonic RC by optical node arrays [5, 13] and photonic time-delay systems [14–17]. For many physical RCs, the inherent dynamics of physical devices are utilized to form random neuron connections of the reservoir. Photonic RC that has rapid signal processing speed and chip-scale integrated capability is one viable solution in developing small size computing unit for edge computing platform. Integrated photonic RC has been realized on an on-chip micro-ring or micro-cavity [18, 19] and interconnected-node mesh reservoir [5, 13]. Photonic delay line-based RC using virtual nodes in a temporal space with simple ring topology was first demonstrated experimentally in an optoelectronic (OE) oscillator configuration by Paquot *et al* [14] and Larger *et al* [20]. Other versions of photonic delay-line RC implementation have also been reported such as all-optical feedback loop [15] and laser cavity with feedback [16]. These reported photonic delay-line RCs are appealing as they can be constructed by off-the-shelf photonic components.

It is well known that the RC network inference result is strongly influenced by the hyperparameters of the RC network. Grid search [17, 21], Bayesian optimization [22–24] and entropy-based hyperparameter optimization have been reported [25]. Grid search for hyperparameter optimization is commonly used to determine the lowest inference error in a global setting, but this approach often suffers from exceedingly long computation time particularly when there are multiple hyperparameters. Comparing to grid search, Bayesian optimization utilizing probability theory can greatly reduce the number of iterations.

Hyperparameter optimization for physical RC poses unique challenge as it requires the tunability and reconfiguration of a physical system. In this work, we report a practical and simple method that utilizes bifurcation diagrams for hyperparameter optimization. We first establish a time evolution equation of the optical oscillator that closely matches the physical setting of the delay-line reservoir. Next, we model the derived evolution equation to determine the dynamic regimes of the RC system on a bifurcation diagram. We find that the low inference errors are generally obtained near bifurcation points [26] when the system is at the edge of stability just before entering the oscillatory state, which allows us to obtain a much smaller range of physical settings to be tuned for RC hyperparameter optimization. We implemented this approach on three benchmark tasks, i.e. nonlinear channel equalization (NCE), nonlinear auto regressive moving average with the 10th order time lag (NARMA10) and Santa Fe time-series prediction tasks.

The paper will be organized as follows. In section 2, general background on time-delay RC network and modeling of our photonic RC implementation are introduced. Detailed derivation of hardware-based time-evolution equation is also provided in this section. Section 3 introduces the intrinsic dynamics of delay-line RC system while section 4 applies the bifurcation-dynamic guided hyperparameter optimization on three RC benchmark tasks.

2. Photonic true-time-delay RC network

The photonic time-delay RC network studied in this work uses a cyclic topology similar to the setup in Paquot *et al* [14]. The input layer signals are sequentially injected to the reservoir layer via a single physical nonlinear node while the signals in the reservoir circulate on a closed delay-line loop. The internal reservoir states are often referred to as the virtual nodes traveling along the time-delay line in a series of optical pulse trains with varying intensity to form a spatial-temporal dynamic system. A good reservoir should have high dimension, nonlinearity and fading memory [27, 28]. The RC network of this work was implemented on a bench-top photonic time-delay system constructed with commercially available optical and electrical components. A schematic of the photonic delay-line RC is sketched in figure 1 and the components used in the RC construction are listed in supplementary materials.

The reservoir in this work is an optical fiber oscillator that operates at the edge of stability. The capacity of the reservoir is denoted by the virtual node number $N = \tau/\theta$, where τ is the round-trip delay time and θ is the node separation in unit of time. In this work, the optical signal is coupled to the reservoir by an arbitrary waveform generator (AWG) at a sampling rate of N/τ. A reservoir connectivity matrix, W_{res} is used to characterize the node connections within the reservoir. A node mismatch k is defined as $k = (\tau - \tau')/\theta$ where τ' denotes the sample hold time for each input signal. For synchronized operation where $\tau = \tau'$, W_{res} is simply an identity matrix. To enrich the reservoir dynamics with increased interactions among virtual nodes, RC can be tuned to asynchronized regime [14] in which W_{res} is a shifted identity matrix. In this work, we utilized a small node mismatch $k = 2$ for NCE and Santa Fe tasks and the W_{res} matrix non-zero diagonal element '1' is shifted by 2 columns to the right as shown below:

$$W_{\text{res}} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 1 & & 0 \\ & & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & & 1 \\ 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & & 0 \end{bmatrix}.$$

The input signal $u(t)$ is acquired externally and then coupled to the physical RC via a sample-hold operation. A piecewise mask function $m(t)$ is multiplied with $u(t)$ to increase the diversity of the projected input data in the reservoir [14]. The discrete levels of $m(t)$ are randomly generated in the range of $[-1, 1]$ with a time duration of θ for each level and raw input data are normalized to $u(t) \in [-1, 1]$. The period of mask function is set equal to the duration time of each $u(t)$ level τ'. In this work, the input layer is

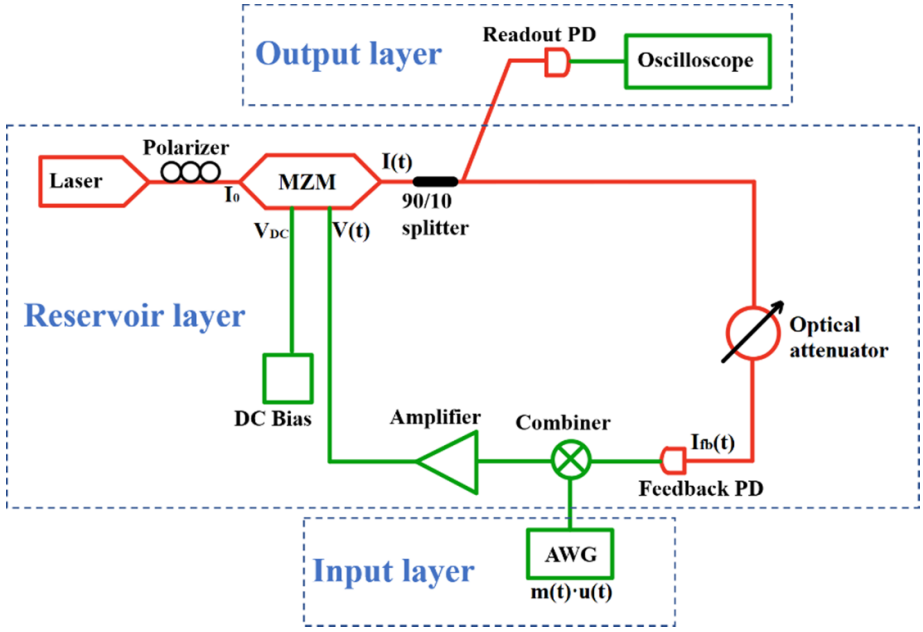

Figure 1. Schematic of photonic implemented time-delay RC system. The optical path is marked in red and the electrical path is marked in green. AWG: arbitrary waveform generator.

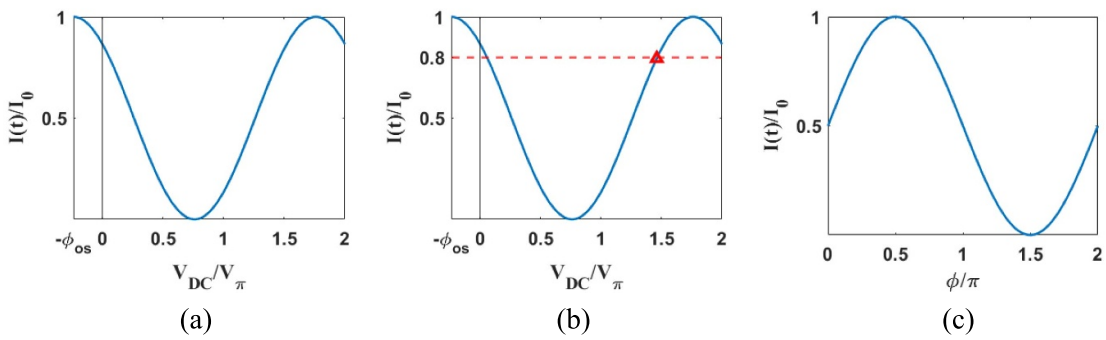

Figure 2. (a) Measured transfer function of the MZM versus external bias. (b) MZM operation near peak. The red triangle marks the DC operating points and red dashed lines represent the reference levels ($x = 0$) of MZM output intensity. Positive x are recorded above reference level and negative x are recorded below reference level. (c) Measured transfer function of the MZM versus hyperparameter φ.

implemented digitally in an external computer and the operation of $m(t) \cdot u(t)$ is carried on in MATLAB program. Duport *et al* [29] reported using cascade Mach–Zehnder modulators (MZMs) to perform the multiplication in optical domain so a fully hardware implementation of the input layer is feasible. The masked input data sequence $m(t) \cdot u(t)$ is fed into the reservoir layer through the AWG. The amplitude of $m(t) \cdot u(t)$ at the AWG output is to be adjusted as a key parameter closely related to β.

In the reservoir layer, a $LiNbO_3$ MZM serves as the single physical node to provide a nonlinear response to the encoded input optical signals. The MZM has a marked half-wave voltage (V_π) of 3.35 V (DC) and 3.7 V (RF). The DC operating point of the MZM is set by an external bias V_{DC}. The measured transfer function of the MZM is plotted in figure 2(a) and a built-in phase offset $\varphi_{os} = 0.24\pi$ is measured at $V_{DC} = 0$. An optical splitter is used at the MZM output to tap out 10% of the optical power to a readout photodetector (PD) in the output layer. An electrical amplifier is used to compensate the optical loss in the delay line as well as provide an easy way to adjust the input gain β and the feedback gain α in addition to the optical attenuator in the delay line loop. The electrical amplifier can be eliminated from the RC hardware setup when the laser power is sufficiently large or the loss in the delay line loop is low. The elimination of the electrical amplifier helps to reduce the power consumption of the RC network as well as reduce thermal noise in the system.

The output intensity $I(t)$ of the MZM with RF port is fed with $V(t)$ can be expressed as

$$I(t) = I_0 \cos^2\left[\frac{1}{2}\left(\frac{\pi V(t)}{V_\pi} + \varphi_0\right)\right] \tag{1}$$

where I_0 is the MZM maximum output power with insertion loss and φ_0 is defined as

$$\varphi_0 = \varphi_{\text{os}} + \frac{\pi V_{\text{DC}}}{V_\pi}. \tag{2}$$

The RF signal $V(t)$ to the MZM comprises the summation of feedback signals after a round trip in the delay line loop and the newly added input $u(t)$ at the combiner and it can be expressed as

$$V(t) = \eta_O \times \eta_{\text{E}} \times \mathcal{R} \times I^*(t-\tau) + \eta_{\text{E}} \times V_{\text{AWG}} \times m(t) \times u(t) \tag{3}$$

where $\eta_O = I_{fb}(t)/I(t)$ is the feedback PD optical power retaining ratio, η_{E} is the total gain/loss of the electrical path (green part in figure 1), $\mathcal{R}$ (unit: V/W) is the multiplication of the feedback PD responsivity with the PD's effective resistance, V_{AWG} is the maximum amplitude of AWG output and $I^*(t-\tau)$ denotes the feedback optical intensity as a function of time.

The output photocurrent (I_{ph}) of both the feedback PD and the readout PD is proportional to the PD input optical intensity while I_{ph} of the feedback PD multiplied with the input impedance of the oscilloscope (50 ohm) yields a voltage that corresponds to the reservoir state value and is displayed at the oscilloscope. For future work of an integrated optical delay-line RC, a transimpedance amplifier can be employed to produce the reservoir states. In Paquot's work [14], the output signal intensity $I(t)$ was normalized following $x(t) = \frac{I(t)-I_0/2}{I_0/2}$ to get $x(t) \in [-1, 1]$ under the assumption that the reference of the output intensity ($x = 0$) is set at $I_{\text{DC}} = I_0/2$ which is different from how the reservoir states are read out from the oscilloscope in our work. Here, we utilized an AC coupled approach to obtain both positive and negative reservoir state. The feedback PD is connected to a high pass filter with a cut-off frequency of 30 kHz so that the DC component of the PD output is blocked. Impedance matched transmission line ensures the AC signals are efficiently delivered to record reservoir states, where a positive/negative AC component of PD output corresponds to positive/negative reservoir state reading.

For illustration, an example is shown in figure 2(b) where V_{DC} is set at 4.9 V, corresponding to a 0.3π phase difference from the peak transmission. This gives a reference of output intensity set at $I_{\text{DC}} = I_0\cos^2(\varphi_0/2) = 0.8I_0$ without any signal fed to the RF port. The instant MZM output $I(t)$ determined by the $V_{\text{DC}} + V(t)$ can then give either a positive (above red line) or a negative (below red line) reservoir state value using I_{DC} as the reference. Normalizing to I_0, the reservoir state can be expressed as

$$x(t) = \frac{I^*(t)}{I_0/2} = \frac{I(t) - I_0\cos^2(\varphi_0/2)}{I_0/2} = \frac{2I(t)}{I_0} - 1 - \cos\varphi_0. \tag{4}$$

Here $x(t)$ varies in $[-1-\cos\varphi_0, 1-\cos\varphi_0]$, effectively having an offset of $\cos\varphi_0$ from $[-1, 1]$ based on V_{DC}. Combining equations (1)–(3) and normalizing all the optical power terms based on equation (4), the time evolution equation can be expressed using hardware parameters as

$$x(t) + \cos\varphi_0 = \cos\left[\frac{\pi\eta_O\eta_{\text{E}}\mathcal{R}I_0}{2V_\pi}x(t-\tau) + \frac{\pi\eta_{\text{E}}V_{\text{AWG}}}{V_\pi}m(t)\,u(t) + \varphi_0\right]. \tag{5}$$

Note the coefficient of $x(t-\tau)$ and $u(t)$ of equation (5) marks the delay-line loop feedback strength α and input signal strength β respectively with the following definition:

$$\alpha = \frac{\pi\eta_O\eta_{\text{E}}\mathcal{R}I_0}{2V_\pi} \tag{6}$$

$$\beta = \frac{\pi\eta_{\text{E}}V_{\text{AWG}}}{V_\pi}. \tag{7}$$

Combining the phase terms yields

$$\varphi = \varphi_0 + \frac{\pi}{2} = \varphi_{\text{offset}} + \frac{\pi V_{\text{DC}}}{V_\pi} + \frac{\pi}{2}. \tag{8}$$

A transfer function of the MZM with respect to the hyperparameter φ defined by equation (8) is plotted in figure 2(c). The general form of a sinusoidal time evolution equation can then be expressed as

$$x(t) + \sin\varphi = \sin[\alpha x(t-\tau) + \beta m(t)\,u(t) + \varphi] \tag{9}$$

where α, β, φ are three hyperparameters of this RC network to be optimized based on the computation tasks. Replacing the time variable t by index n, a discretized time evolution equation can be derived for convenience of theoretical studies:

$$x_i(n) + \sin\varphi = \sin[\alpha x_i(n-1) + \beta m_i u(n) + \varphi] \tag{10}$$

where n represents the nth input data and i denotes the ith virtual node on the delay line.

In the output layer, a one-step training is performed by calculating the output matrix W_{out} via ridge regression

$$W_{\text{out}} = YX^T\left(XX^T + \lambda I\right)^{-1} \tag{11}$$

where Y is the targeted output matrix consisting of targeted outputs $y(n)$, and X is the reservoir states matrix concatenating the projected reservoir state vectors of each input data point. The superscript T represents the matrix transpose operation and I denotes the identity matrix. λ is a regularization parameter to prevent overfitting and is chosen according to the virtual node number N' and the computation task.

3. Delay-line RC system dynamics

The dynamics of OE time-delay systems have been investigated in several works [30–32]. Passive optical systems subject to delayed feedback can be described by Ikeda delay differential equation (DDE) or slightly modified forms [33] to model the nonlinearity [34]. The dynamics of the delay-line RC used in this work can be described as

$$\tau_r x'(t) = -x(t) + \sin[\alpha x(t-\tau) + \varphi] - \sin\varphi \tag{12}$$

where the dimensionless dynamical variable x represents the normalized MZM output power and α, φ follows the same notation as discussed in section 2. Though the detailed form of the DDEs may vary due to system configuration variations, two parameters are common for OE delayed systems in characterizing the dynamics: relaxation time τ_{r} and delay time τ. The parameter τ_r in equation (12) characterizes how fast the system reacts to the input changes and is limited by the system cutoff frequency f_{H}. In our RC network, the slowest component is the feedback PD (and the readout PD) with a marked bandwidth of 1 GHz. The estimated τ_r of the RC is $\tau_{\text{r}} = 1/2\pi f_{\text{H}} \approx 0.16\,\text{ns}$ [16] which is several orders of magnitude smaller than the oscillator cavity delay time τ. For the case of $\tau_{\text{r}}/\tau \ll 1$, the derivative term $x'(t)$ in equation (12) can be neglected as suggested in [33]. In some delayed dynamical systems [8, 10], τ_{r} is comparable to the node separation θ to incite interactions between the neighboring nodes to further boost the reservoir dynamics. In this work, we set $\tau_{\text{r}} \ll \theta$ and $\tau_{\text{r}} \ll \tau$ so neighbor nodes interaction is negligible. Consider a discrete-time system with index n, equation (12) can then be simplified to

$$x(n) = \sin[\alpha x(n-1) + \varphi] - \sin\varphi. \tag{13}$$

Equation (13) shows an explicit mapping of the current dynamical state $x(n)$ with the past state $x(n-1)$, which agrees with the time evolution equation (10) when $\beta = 0$. The stability of the system can be revealed from a Hopf bifurcation diagram. In general, the system is stable for small α and oscillatory or chaotic when α value is large.

We define α_{H} as the α value at the onset of oscillatory state on the Hopf bifurcation diagram. When the system has no input feed, i.e. $\beta = 0$, varying φ from 0 to 2π, α_H versus φ is plotted in figure 3 given the initial state $x(0) = 1$. Above the α_H curve, the system is oscillatory or chaotic, and below it, the system is stable. The reservoir layer, utilizing the system dynamics, projects the input signals to a higher dimensional space to increase separation ability of the inputs signals. Dimensionality expansion is the underline physics of RC inference. Richer dynamics of the reservoir gives better RC inference results. Hence, the hyperparameters should be chosen so that the RC operates near the bifurcation point where the system has the richest dynamics while still in the stable regime. Below the $\alpha_{\text{H}}(\varphi)$ curve, we marked a region in cyan bounded by $\alpha_{\text{H}}(\varphi) - \Delta$ (red dash line) where Δ is a small value to visualize a set of ideal hyperparameter combinations on the $\varphi - \alpha$ plane.

The $\alpha_{\text{H}}(\varphi)$ curve in figure 3 reveals the intrinsic dynamics of the system without input signals, i.e. $\beta = 0$. As shown in equations (9) and (10), the input signal $m(t) \cdot u(t)$ multiplied with the input gain β produces an extra phase to the MZM cosine transfer function. When the input data is normalized to $[-1, 1]$, the input layer signals, i.e. $m(t) \cdot u(t)$ projected to the reservoir layer at the combiner is proportional to [-β, β]. If β is too small, the extra phase induced by β will cause the reservoir states to only vary in a small portion of the

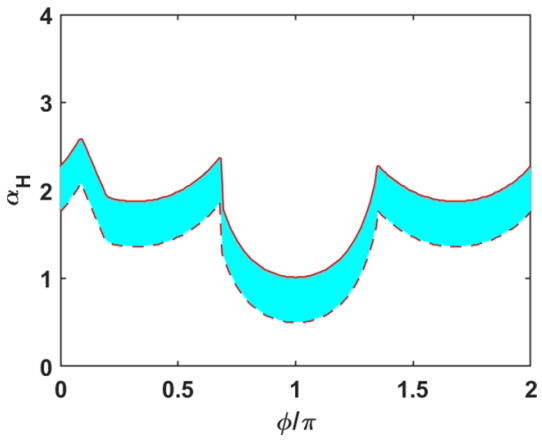

Figure 3. Simulated values of Hopf bifurcation point α_H where the system goes from steady states to oscillation as the function of phase φ. Initial state $x(0) = 1$. Dashed red curves are plotted by shifting down the simulated curves (solid) by 0.5, i.e. $\alpha_H(\varphi) - 0.5$.

sinusoidally varying curve of the nonlinear node. For example, when the MZM is biased near its quadrature point, we can only obtain a near linear response of the reservoir with small β, so the reservoir layer acts more like a linear classifier rather than a complex dynamic system. On the other hand, if β is large, the feedback signals produced at the PD become insignificant compared to the input at the combiner, which diminishes the memory effect of the RC. β is mainly tuned by adjusting the amplitude of the AWG output and in general has minor impact on the shape of optimized regions on $\varphi - \alpha$ plane. However, proper choice of β can give us lower error rates for different RC tasks and more relaxed restriction to α and φ conditions.

4. Benchmark tasks using bifurcation dynamics guided hyperparameter optimization

For the photonic delay-line RC of this study, the feedback gain (α), input gain (β) and bias phase offset of the modulator (φ) are the RC hyperparameters to be adjusted, though the RC inference results are not equally sensitive to all these hyperparameters. The feedback gain and the phase value are more crucial in determining the RC network performance. Based on the theoretical study of $\alpha - \varphi$ dynamic regime while choosing the proper input gain β, the hyperparameter optimization can be simplified to a 2D grid search. Furthermore, using the simulated bifurcation diagram based on hardware parameters extracted to construct the evolution equation, there is no need to run a full space 2D optimization of α and φ while a much reduced $\alpha - \varphi$ grid range would produce good results. This bifurcation dynamic guided optimization utilizing mathematic modeling of hardware RC leads to a significant time reduction in hardware setting optimization.

We tested our method of bifurcation dynamic guided hyperparameter optimization on three commonly studied benchmark tasks: NCE, NARMA10 and Santa Fe tasks. For the RC numerical study based on equation (10), we perform a fine 2D grid search for α and φ in which φ and α are scanned in a range of 0 to 2π with a step of 0.02π and 0 to 4 with a step of 0.02, respectively. This mathematic modeling is carried on in MATLAB with fine resolution to achieve the best RC inference results. β is set at different fixed values for three benchmark tasks and more discussion of β effect on RC performance will be provided in section 4(d).

To examine how closely the experimental testing results match the simulated results, we also run a 2D grid search for α and φ configurations on the photonic delay line RC. The hyperparameter φ, tuned by changing the value of MZM DC bias voltage, is tested with a step of $\pi/6$ for each setting. Hyperparameter α and β are experimentally adjusted by tuning the optical attenuator and AWG maximum output amplitude, respectively. For the photonic RC, the reservoir enters either an oscillatory or chaotic state at large α so the photonic RC hardware settings corresponding to large α are excluded from the experimental tests. For very small α value, close to zero, the photonic RC performance is limited by the system noise originating from analog-to-digital conversion (ADC) of oscilloscope, laser relaxation oscillations, and PD dark current and shot noise, etc [17, 32], so α with small values (< 0.05) are also excluded from the tests. In general, we choose α with a small interval (~ 0.05) near chaotic states and a larger interval as α decreases. The total number of tested (φ, α) settings is $80 - 100$ for each task, sufficient to distinguish the error rate difference caused by varying the hyperparameter setting. Detailed experimental settings of three tasks are listed in table 1.

For our current experimental setup, optical attenuator and DC bias voltage are tuned manually, making it time consuming to adjust the hardware for all possible combinations of φ and α values, therefore a much larger grid size is utilized compared with numerical simulation. The theoretical modeling results are plotted

Table 1. Detailed experimental settings of NCE, NARMA10 and Santa Fe tasks.

Benchmark RC task	Fiber length	Round-trip delay time	AWG processing speed	θ	N	N'	β	λ
NCE	125 m	650.6 ns	250 MSa s^{-1}	4 ns	163	161	0.74	0.001
NARMA10	1 km	4950 ns	125 MSa s^{-1}	8 ns	619	613	1.07	0.01
Santa Fe	125 m	650.6 ns	250MSa s^{-1}	4 ns	163	161	1.75	0.001

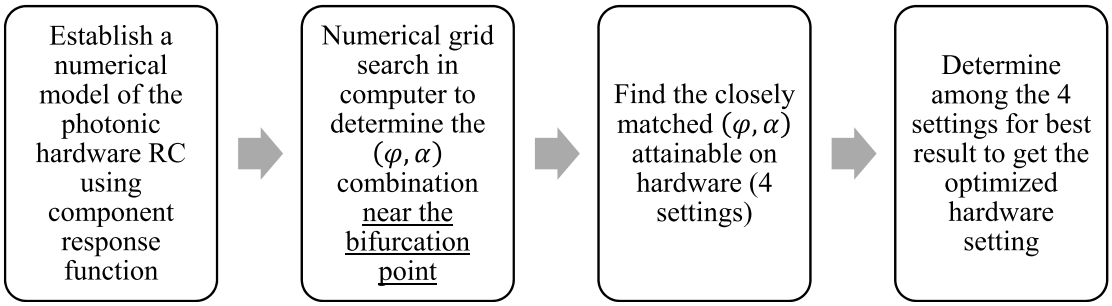

Figure 4. Flow diagram showing the overall optimization process.

in 2D color contour maps where 'islands' of good inference results are identified. Based on the optimal values of φ and α obtained by theoretical modeling, we identify four combinations of (φ, α) in the photonic delay-line RC setting that is closest to simulation. An empirical factor 0.6 is applied in determining the optimal α and scaled in the 2D plots of the experimental results. The overall optimization process is shown through the flow diagram in figure 4.

4.1. NCE task

NCE task is widely studied in the RC community [14, 15, 35, 36]. The input $d(n)$ is randomly generated from the set {-3, −1, 1, 3} following uniform distribution. The nonlinear channel provides a transformation following

$$q(n) = 0.08d(n+2) - 0.12d(n+1) + d(n) + 0.18d(n-1) - 0.1d(n-2) + 0.091d(n-3) - 0.05d(n-4) + 0.04d(n-5) + 0.03d(n-6) + 0.01d(n-7)$$

$$u(n) = q(n) + 0.036q(n)^2 - 0.011q(n)^3 + v(n) \tag{14}$$

where $v(n)$ is the i.i.d. Gaussian noise term with zero mean. In this work, the signal-to-noise ratio is set to be 28 dB and a dataset length of 3000, a training length of 1200, and a testing length of 1600 is used.

The symbol error rate (SER) of the 2D RC simulated result is plotted in a color contour map with varying α and φ, shown in figure 5(a). Each contour line corresponds to an SER reading while the filled-in colored area represents the level of SER. The SER contour plot is close to symmetric about $\varphi = \pi$. SER lower than 0.01 is widely achieved over a large φ range near two quadrature points for $\alpha \in [0.25, 1]$. SER lower than 0.001 is found to form a small 'island' near $\varphi = \pi$ and a larger 'island' near $\varphi = 2\pi$ (0), both marked in dark blue. Numerical modeling of the RC gives SER $< 6 \times 10^{-4}$ at $\varphi = 1.82\pi$, $\alpha = 0.64$ marked by red 'x' mark in figure 5(a).

The $\alpha - \varphi$ grid search was performed on the photonic delay-line RC and the SER results are shown in figure 5(b). Two areas of SER $< 6 \times 10^{-4}$ are identified in the experimental grid search: one is between $\varphi = 0$ and $\varphi = 0.5\pi$ and the other one between $\varphi = 1.5\pi$ and $\varphi = 2\pi$. The achieved SER is on the same order as Paquot *et al* [14] that reported SER of 1.3×10^{-4} and we expect the upper bound of our best SER to be smaller if more data points are utilized for testing. These low SER areas in the photonic RC testing are closely matched with the RC numerical modeling of equation (10). Due to φ being continuous with a period of 2π, these two areas are indeed connected and correspond to the quadrature point $\varphi = 0$ (2π). The general trend of the SER by photonic RC matches well with the RC by numerical modeling. Four closest grid points to the optimal (φ, α) were identified, marked by green 'x' marks in figure 5(b) on the photonic RC 2D SER map and summarized in table 2.

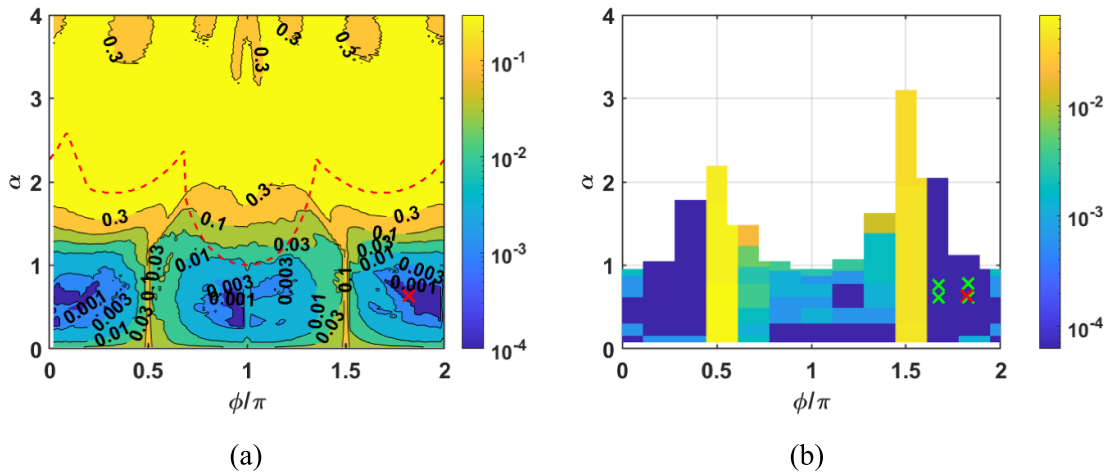

Figure 5. NCE 2D SER results: (a) RC theoretical modeling. The red and blue dashed curves correspond to $\alpha_{\mathrm{H}} - \varphi$ relation in figure 3, respectively. The red 'x' mark is the setting giving rise to the lowest SER. (b) Time-delay photonic RC implementation. Red 'x' mark represents the optimal (φ, α) condition suggested by numerical study and four closest photonic RC hardware settings are marked by green 'x' marks.

Table 2. Matched photonic hyperparameter settings for NCE task based on the numerical study.

	φ	α	SER
Test setting 1	1.83π	0.62	0
Test setting 2	1.83π	0.78	0
Test setting 3	1.67π	0.62	6.25×10^{-4}
Test setting 4	1.67π	0.77	0

To examine how the numerical RC matches with the photonic RC, we performed 2D α—φ grid search by both numerical and experimental approach. In practice, the photonic RC hyperparameter optimization can be reduced to examining those four (φ, α) settings closely matching the bifurcation dynamics analysis to set the hardware parameters of the photonic RC. This would allow significant time reduction in tedious hardware adjusting.

4.2. NARMA10 task

The emulation of the nonlinear auto regressive moving average model, i.e. NARMA10 task, is a commonly used benchmark prediction task for nonlinear dynamic systems. It was first introduced in [1] and studied as a benchmark prediction task in [5, 8]. The discrete time input $u(n)$ and targeted output $y(n)$ follows the dynamic

$$y(n+1) = 0.3y(n) + 0.05y(n) \cdot \sum_{i=0}^{9} y(n-i) + 1.5u(n)\,u(n-9) + 0.1 \tag{15}$$

where $u(n)$ follows a uniform distribution over the range of [0, 0.5] and is randomly generated. The goal of this task is to predict the output $y(n)$ given the input $u(n)$. A dataset with a length of 3000 is used to run the task, with a training length of 1000, testing length of 1800 and first 200 datapoints discarded. The RC performance of NARMA10 task is characterized by the normalized mean square error (NMSE) [29]. The color map of NMSE with respect to φ, α by numerically modeled RC is shown in figure 6(a). There are five 'islands' that correspond to low NMSE on the bifurcation diagram. A large 'island' of ideal (φ, α) combinations was found near $\varphi = \pi$, just below the system bifurcation, on which the lowest NMSE by the RC numerical simulator is found to be 0.0256 at $\varphi = 0.88\pi$, $\alpha = 0.92$, marked by red 'x' mark as shown in figure 6(a). Those two isolated dark blue islands near $\alpha = 2$ are not experimentally measurable due to chaotic response of the photonic RC system. Those two low NMSE areas at bottom left and right of the bifurcation diagram ($\alpha \sim 1$), near symmetric and connected at $\varphi = 0$ and 2π considering an unfolded diagram, are also ideal operation regimes for NARMA10 task. If we consider NMSE $\leqslant 0.2$ acceptable, both lower 'islands' near two quadrature points with $\alpha \sim 1$ meet the criterion while the 2D photonic grid search matches well with the numerical RC modeling.

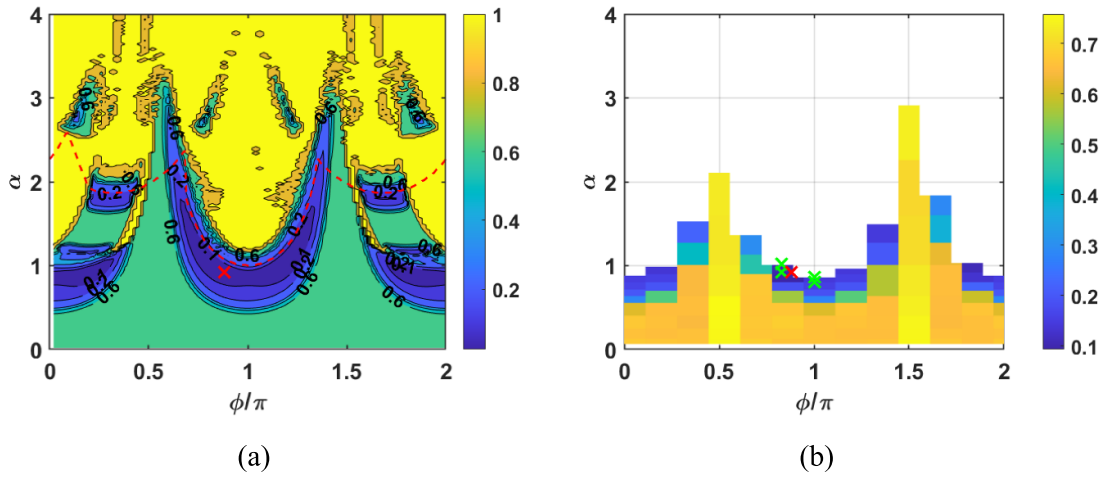

Figure 6. NARMA10 2D NMSE results: (a) RC theoretical modeling. The red and cyan dashed curves correspond to $\alpha_H - \varphi$ relation in figure 3, respectively. The red 'x' mark indicates the condition with the lowest NMSE. (b) Time-delay photonic RC implementation. Four closest grid points by green 'x' marks are labeled near the optimal (φ, α) by numerical RC simulator.

Table 3. Matched photonic hyperparameter settings for NARMA10 task based on the numerical study.

	φ	α	NMSE
Test setting 1	0.83π	0.92	0.101
Test setting 2	0.83π	1.02	0.0930
Test setting 3	π	0.87	0.142
Test setting 4	π	0.80	0.155

The color map trends of experimental results shown in figure 6(b) overall matches the numerical RC simulator results, where the low NMSE [dark blue in figure 6(b)] area is also found near two quadrature points and close to bifurcation. Those 4 closet grid points on the photonic RC to the optimal (φ, α) by numerical RC simulator were marked by green 'x' marks on Fib. 5(b) the calculated NMSE results are summarized in table 3. Using photonic RC, the lowest NMSE is 0.0930 corresponding to setting 2 on table 3, better than Paquot's [14] result NMSE = 0.168 obtained from similar experimental setup as ours. Photonic RC testing overall yields higher prediction error rate than the RC simulator mainly because of the quantization noise in the ADC process at the output layer [17]. Both simulated and experimental results show that peak and valley of MZM transmission should be avoided.

4.3. Santa Fe task

The Santa Fe laser time-series prediction task is another well studied RC benchmark task [9, 18]. We performed one-step ahead prediction using a dataset with 4100 data points where the first 100 data points are discarded, the next 3000 are used for training, and the remaining 1000 data points are used for testing.

The 2D NMSE plots by numerical RC simulator and photonic RC are plotted in figure 7, both indicating higher prediction error near 0.5π and 1.5π at low α values. Compared to the NCE and NARMA10 tasks, Santa Fe task is less sensitive to hyperparameter settings as suggested by both numerical RC simulator and the photonic RC. The simulated contour map in figure 7(a) shows a band of all φ values near $\alpha \sim 0.8$ that gives NMSE < 0.01. Simulation reveals a lowest NMSE of 0.0054 at $\varphi = 1.36\pi$, $\alpha = 1.1$. Four closely matched hyperparameter settings on the photonic RC are identified, two of which correspond to the modulator transmission valley so not considered. The other two matched settings are summarized in table 4. Test setting 1 gives a NMSE of 0.0341, which is 6.3 times higher than the RC simulator result mainly due to the quantization noise similar as NARMA10 task. Both settings result in lower NMSE than Larger's result of NMSE = 0.124 that used similar photonic RC setup.

4.4. Hyperparameter β effect

In this work, we performed a simple analysis of the effect of β to the RC system on the numerical RC simulator. We scanned β from 0.5 to 2.5 with a relatively large step of 0.1. The choice of lower end of β is due to the concern of photonic RC noise, while the AWG output amplitude level determines the highest β. A 3D grid hyperparameter scan with lower resolution of α and φ is performed and the lowest SER or NMSE varying with β for those three benchmark tasks are plotted in figure 8. In the tested range of β, the NMSE

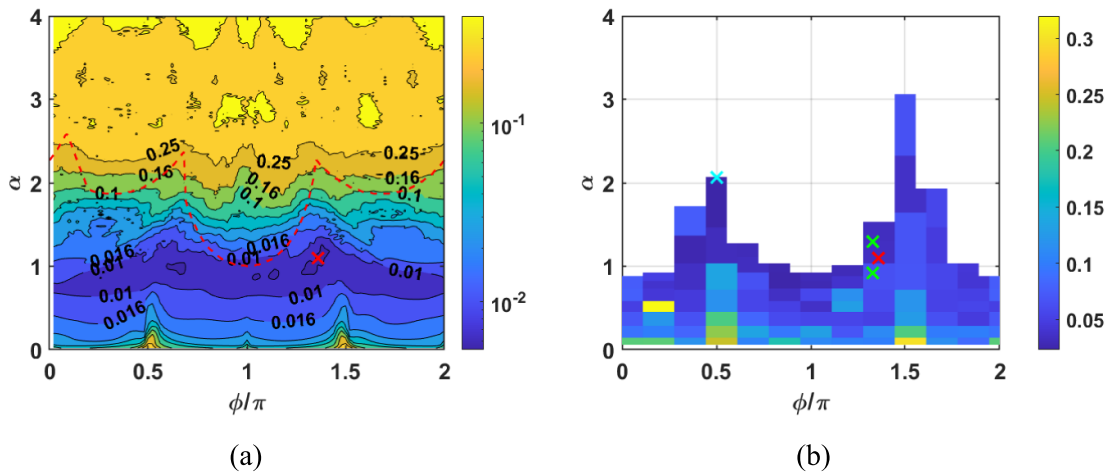

Figure 7. Santa Fe 2D NMSE results: (a) RC theoretical modeling. The red and yellow dashed curves correspond to $\alpha_H - \varphi$ relation in figure 3, respectively. The red 'x' mark corresponds to the condition leading to the lowest NMSE. (b) Time-delay photonic RC implementation. Red crossing represents the optimal (φ, α) condition suggested by numerical study and two closest photonic RC testing conditions (excluding MZM valley transmission) are marked with green 'x' marks. The cyan 'x' mark represents the condition giving lowest NMSE of 0.0232 in photonic RC testing.

Table 4. Matched photonic hyperparameter settings for Santa Fe task based on the numerical study.

	φ	α	NMSE
Test setting 1	1.33π	1.30	0.0341
Test setting 2	1.33π	0.92	0.0648

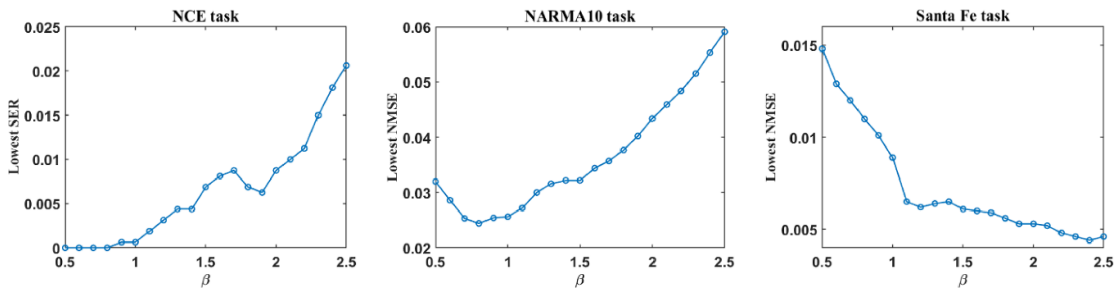

Figure 8. Simulated lowest error rate (SER/NMSE) as the function of β for three benchmark tasks: (a) NCE task; (b) NARMA10 task; (c) Santa Fe task.

only changes for a factor 2.4 and 3 for the NARMA 10 task and Santa Fe task, respectively. As a comparison, NMSE can vary several orders of magnitude with (φ, α) choices.

For the NCE task, SER $< 10^{-3}$ is obtained for $\beta < 1$. In our photonic RC test, $\beta = 0.7$ was set that gives SER $= 0$. For the NARMA10 task, a valley of low NMSE values is obtained for β in the range of 0.7–1.1. In the photonic RC test, $\beta = 1$ was selected as larger β is preferred for relatively higher signal-to-noise ratio. For the Santa Fe task, the NMSE decreases monotonically with β that shows a different trend as the NCE and NARMA 10 task. During the Santa Fe task test, due to an equipment failure, we changed to a different AWG (Tektronix 70 000) that has a maximum output voltage $V_{AWG} = 0.25V$, corresponding to $\beta = 1.8$. We expect reduced NMSE for the Santa Fe task if a higher AWG output voltage is available.

5. Conclusion

In this work, we explored a practical engineering approach in selecting proper hyperparameter values of a photonic delay-line RC through bifurcation dynamic analysis. The time-evolution equation of the photonic RC is derived and modeled numerically based on the hardware settings. The intrinsic dynamics and bifurcation behaviors of the time-delay system are analyzed to give insight on how each hyperparameter affects the RC dynamics, so a reduced range of grid search is possible in determining the optimal operation conditions of a RC. It shows that a good RC performance can be obtained by tuning the delay-line system

close to the bifurcation point while operating in the steady-state regime. The low-error-rate region shows good alignment with the $\alpha_{\mathrm{H}}(\varphi)$ curve near quadrature points in both the numerical study and the experimental hardware testing for all three benchmark tasks. A general good agreement in the 2D bifurcation diagram is obtained between the numerical model and the photonic RC experimental test. Among the three hyperparameters of the photonic RC, delay-line loop feedback strength α, modulator phase φ, and input signal strength β, the photonic RC appears to be more sensitive to the choice of α and φ so a 2D grid optimization for α and φ was performed at a fixed β value chosen from both empirical as well as a simple numerical analysis. Using the numerical RC modeling results, 4 experimental settings of (α, φ) are exanimated on the photonic RC and assessed by NMSE or SER. Good agreement is obtained between the photonic RC test results and the numerical simulation.

Data availability statement

The data cannot be made publicly available upon publication because they are not available in a format that is sufficiently accessible or reusable by other researchers. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

The experimental components are partially supported by AFRL under the Education Partnership Agreement of EPA # 13-RI-EP-34.

ORCID iDs

Meng Zhang https://orcid.org/0000-0001-7182-8110
Zhizhuo Liang https://orcid.org/0000-0002-7023-1619
Z Rena Huang https://orcid.org/0000-0002-0667-903X

References

[1] Jaeger H 2002 Adaptive nonlinear system identification with echo state networks *Advances in Neural Information Processing Systems* vol 15
[2] Verstraeten D, Schrauwen B, Stroobandt D and Van Campenhout J 2005 Isolated word recognition with the liquid state machine: a case study *Inf. Process. Lett.* **95** 521–8
[3] Jalalvand A, Van Wallendael G and Van de Walle R 2015 Real-time reservoir computing network-based systems for detection tasks on visual contents *2015 7th Int. Conf. on Computational Intelligence, Communication Systems and Networks* (IEEE)
[4] Soh H and Demiris Y2012 Iterative temporal learning and prediction with the sparse online echo state Gaussian process *The 2012 Int. Joint Conf. on Neural Networks (IJCNN)* (IEEE)
[5] Vandoorne K *et al* 2014 Experimental demonstration of reservoir computing on a silicon photonics chip *Nat. Commun.* **5** 1–6
[6] Bauduin M *et al* 2015 Equalization of the non-linear satellite communication channel with an echo state network *2015 IEEE 81st Vehicular Technology Conf. (VTC Spring)* (IEEE)
[7] Argyris A, Bueno J and Fischer I 2018 Photonic machine learning implementation for signal recovery in optical communications *Sci. Rep.* **8** 1–13
[8] Appeltant L, Soriano M C, Van der Sande G, Danckaert J, Massar S, Dambre J, Schrauwen B, Mirasso C R and Fischer I 2011 Information processing using a single dynamical node as complex system *Nat. Commun.* **2** 1–6
[9] Soriano M C, Ortin S, Keuninckx L, Appeltant L, Danckaert J, Pesquera L and van der Sande G 2014 Delay-based reservoir computing: noise effects in a combined analog and digital implementation *IEEE Trans. Neural Netw. Learn. Syst.* **26** 388–93
[10] Haynes N D *et al* 2015 Reservoir computing with a single time-delay autonomous Boolean node *Phys. Rev.* E **91** 020801
[11] Torrejon J *et al* 2017 Neuromorphic computing with nanoscale spintronic oscillators *Nature* **547** 428–31
[12] Nakane R, Tanaka G and Hirose A 2018 Reservoir computing with spin waves excited in a garnet film *IEEE Access* **6** 4462–9
[13] Katumba A *et al* 2017 A multiple-input strategy to efficient integrated photonic reservoir computing *Cogn. Comput.* **9** 307–14
[14] Paquot Y, Duport F, Smerieri A, Dambre J, Schrauwen B, Haelterman M and Massar S 2012 Optoelectronic reservoir computing *Sci. Rep.* **2** 1–6
[15] Duport F, Schneider B, Smerieri A, Haelterman M and Massar S 2012 All-optical reservoir computing *Opt. Express* **20** 22783–95
[16] Brunner D, Soriano M C, Mirasso C R and Fischer I 2013 Parallel photonic information processing at gigabyte per second data rates using transient states *Nat. Commun.* **4** 1–7
[17] Soriano M C, Ortín S, Brunner D, Larger L, Mirasso C R, Fischer I and Pesquera L 2013 Optoelectronic reservoir computing: tackling noise-induced performance degradation *Opt. Express* **21** 12–20
[18] Zhang H, Feng X, Li B, Wang Y, Cui K, Liu F, Dou W and Huang Y 2014 Integrated photonic reservoir computing based on hierarchical time-multiplexing structure *Opt. Express* **22** 31356–70
[19] Sunada S and Uchida A 2019 Photonic reservoir computing based on nonlinear wave dynamics at microscale *Sci. Rep.* **9** 1–10
[20] Larger L, Soriano M C, Brunner D, Appeltant L, Gutierrez J M, Pesquera L, Mirasso C R and Fischer I 2012 Photonic information processing beyond Turing: an optoelectronic implementation of reservoir computing *Opt. Express* **20** 3241–9
[21] Bueno J, Brunner D, Soriano M C and Fischer I 2017 Conditions for reservoir computing performance using semiconductor lasers with delayed optical feedback *Opt. Express* **25** 2401–12
[22] Yperman J and Becker T 2016 Bayesian optimization of hyper-parameters in reservoir computing (arXiv:1611.05193)

[23] Aaron G, Pomerance A and Gauthier D J 2019 Forecasting chaotic systems with very low connectivity reservoir computers *Chaos* **29** 123108
[24] Antonik P *et al* 2021 Bayesian optimisation of large-scale photonic reservoir computers *Cogn. Comput.* 1–9
[25] Carroll T L 2021 Optimizing reservoir computers for signal classification *Front. Physiol.* **12** 893
[26] Goldmann M, Köster F, Lüdge K and Yanchuk S 2020 Deep time-delay reservoir computing: dynamics and memory capacity *Chaos* **30** 093124
[27] Maass W, Natschläger T and Markram H 2002 Real-time computing without stable states: a new framework for neural computation based on perturbations *Neural Comput.* **14** 2531–60
[28] Jaeger H 2002 Short term memory in echo state networks. GMD-report 152 *GMD-German National Research Institute for Computer Science* (available at: www.faculty.jacobs-university.de/hjaeger/pubs/STMEchoStatesTechRep.pdf.)
[29] Duport F, Smerieri A, Akrout A, Haelterman M and Massar S 2016 Fully analogue photonic reservoir computer *Sci. Rep.* **6** 1–12
[30] Hopf F A, Kaplan D L, Gibbs H M and Shoemaker R L 1982 Bifurcations to chaos in optical bistability *Phys. Rev.* A **25** 2172
[31] Kouomou Y C, Colet P, Larger L and Gastaud N 2005 Chaotic breathers in delayed electro-optical systems *Phys. Rev. Lett.* **95** 203903
[32] Callan K E, Illing L, Gao Z, Gauthier D J and Schöll E 2010 Broadband chaos generated by an optoelectronic oscillator *Phys. Rev. Lett.* **104** 113901
[33] Erneux T, Larger L, Lee M W and Goedgebuer J-P 2004 Ikeda Hopf bifurcation revisited *Physica* D **194** 49–64
[34] Ikeda K 1979 Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system *Opt. Commun.* **30** 257–61
[35] Vinckier Q, Duport F, Smerieri A, Vandoorne K, Bienstman P, Haelterman M and Massar S 2015 High-performance photonic reservoir computer based on a coherently driven passive cavity *Optica* **2** 438–46
[36] McDonald N, Hollingsworth A, Beninati S, Lim C, Ji W, Shi C, Agiorgousis M L, Zhang S and Huang Z R 2020 Analysis of an ultra-short true time delay line optical reservoir computer *J. Lightwave Technol.* **38** 3584–91