Coherent all Optical Reservoir Computing for Equalization of Impairments in Coherent Fiber Optic Communication Systems

Shiva Kumar, *Member, IEEE*, Mahmoud M. T. Maghrabi, Mohamed H. Bakr, *Senior Member, IEEE*, Toshihiko Hirooka, *Member, IEEE*, and Masataka Nakazawa, *Life Fellow, IEEE*

***Abstract*—Reservoir computing (RC) is a bio-inspired framework suited for temporal data processing. Here we propose an all optical coherent RC for the equalization of the impairments of fiber optic (FO) system based on coherent detection. We compare the performances of the RC systems based on semiconductor saturable absorber mirror (SESAM) and highly nonlinear waveguide (HNLW). For a dispersion managed long haul FO system in which the nonlinear penalty is dominant, we find that the SESAM outperforms the HNLW due to its larger nonlinear memory. In contrast, for the equalization of dispersive impairments in a short haul FO system such as data center networks, we find that the HNLW has a much superior performance due to its larger linear memory.**

***Index Terms*—Dispersion, fiber optics, fiber nonlinear effects, optical communication, reservoir computing, recurrent neural network.**

I. Introduction

THE reservoir computing (RC) is a class of recurrent neural network (RNN) introduced by Jaeger [1] and Maass et al. [2]. Several experimental demonstrations of the RC concept have been carried out, including electronic RC with a nonlinear node and delayed feedback [3], optoelectronic RC [4], [5], [6], and all-optical RC [7], [8], [9], [10], [11], [12], [13], [14]. The all-optical RC was first reported by Duport et al. [7]. It was based on a delayed feedback dynamical system and the nonlinearity was provided by the gain saturation of semiconductor optical amplifier (SOA). Since then, the all-optical RC based on semiconductor laser with feedback [8], micro ring resonators [9], and semiconductor saturable absorber mirror (SESAM) [10] have been reported. Nakajima et al. experimentally demonstrated the nonlinear equalization of wireless communication channel using the photonic complex-valued RC [11]. A passive photonic chip based on RC has been demonstrated by Vandoorne et al. [12]. Vinckier et al. have demonstrated the photonic implementation of a RC system based on a coherently driven passive fiber cavity processing analog signals [13]. Argyris et al. introduced a real-valued RC scheme for data classification of severely distorted optical signal in a fiber optic communication system based on direct detection and demonstrated experimentally an improvement in bit error rate (BER) by two orders of magnitude [14]. Equalization of linear and nonlinear impairments of fiber optic system based on direct detection using the RC has drawn significant attention recently [15], [16], [17], [18], [19], [20] owing to its ease of training.

The RC consists of three layers: the input, the reservoir and the output. The reservoir is a recurrent network with randomly connected nonlinear nodes. The input signal vector after multiplication by a fixed random weight matrix is connected to the reservoir nodes (see Fig. 1). A single nonlinear node with delayed feedback is sufficient to emulate a recurrent neural network [3]. The output signal vector is obtained by multiplying the states of reservoir nodes by an output weight matrix. The advantage of the RC over the conventional feedforward artificial neural network (ANN) is that only the output weights need to be trained which significantly reduces the training effort.

In this paper, a coherent all-optical RC is proposed for the equalization of the linear and nonlinear impairments of a coherent fiber optic communication system. There are major differences between the real valued RC for the system based on direct detection and the coherent RC for the system based on coherent detection, for the equalization of the impairments of a fiber optic system. First, in the case of direct detection (DD), the receiver does not have access to the optical phase. The real-valued RC following the DD (or any equalizer following the DD) cannot fully compensate for dispersive and nonlinear impairments due to the fiber optic link since the optical phase information is lost during the detection process. In contrast, the coherent RC placed just before the coherent detection can, in principle, exactly compensate for these impairments. Second, inputs of the real-valued RC are in electrical/digital domain and hence, the overall processing speed is limited by electronics. The inputs of the coherent RC, however, are in optical domain and therefore, we can take advantage of the fast all-optical processing techniques [21]. Third, typically the optical RC schemes proposed for the equalization

Received 2 September 2024; accepted 12 September 2024. Date of publication 17 September 2024; date of current version 30 September 2024. This work was supported in part by the Japan Society for Promotion of Science (JSPS) invitation fellowship, under Grant JSPS KAKENHI and in part by Scientific Research under Grant 24K00878, and in part by Natural Sciences and Engineering Research Council (NSERC), Canada discovery grants for this work. *(Corresponding author: Shiva Kumar.)*

Shiva Kumar, Mahmoud M. T. Maghrabi, and Mohamed H. Bakr are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4L8, Canada (e-mail: skumar@mcmaster.ca).

Toshihiko Hirooka and Masataka Nakazawa are with the Research Institute of Electrical Communication, Tohoku University, Sendai 980-8577, Japan.

Digital Object Identifier 10.1109/JPHOT.2024.3462438

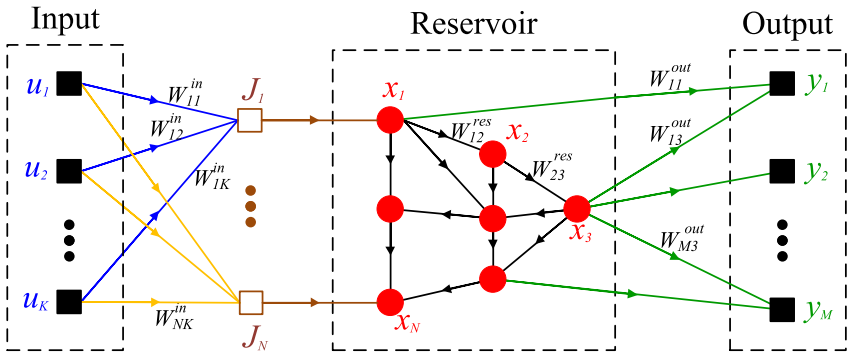


Fig. 1. Schematic of a traditional RC system.

of fiber impairments require back-to-back optical-to-electronic (O/E) conversion (at the optical receiver of the communication system) and electronic-to-optical (E/O) conversion (at the input of an optical RC) whereas in the proposed scheme, the signal to be equalized is already in the optical domain. There is no need for E/O conversion. Our results show that both dispersive and nonlinear impairments of a fiber optic communication system can be significantly compensated for by placing the coherent all-optical RC just before the coherent detection.

Staffoli et al. have developed a silicon based photonic neural network (PNN) device to compensate for chromatic dispersion in fiber optic communication system based on the direct detection [22]. The PNN device is introduced after the fiber optic link just before the receiver. The PNN device consists of a multi-mode interferometer for splitting the signal field into four waveguides, spirals for introducing the delay in each waveguide and phase shifters realized with current controlled heaters. Thus, the signal in each waveguide is delayed depending on the length of the spiral and multiplied by a complex weight that depends on the loss of the spiral and the phase shift. Experimental results showed a significant improvement in the BER demonstrating the ability of the PNN device to equalize the chromatic dispersion. In this paper, we propose to use the delay and multiplication by a complex weight in each waveguide quite similar to [22] for the input and output of the reservoir. In addition, we have a reservoir loop consisting of a nonlinear device such as semiconductor saturable absorber mirror (SESAM) or highly nonlinear waveguide (HNLW) acting as a nonlinear activation function.

Sorokina et al. have proposed a fiber echo state network for dual-quadrature signal processing [23]. Using the Kerr nonlinearity in fiber, they have achieved the sine function as the nonlinear activation function. The input weights and reservoir weights are implemented in electrical domain and hence, the processing speed is limited by the electronic bottleneck. Sorokina extended the work of [23] by replacing the electrical input and reservoir weights with the optical weights realized using the dispersive and nonlinear properties of the fiber [24]. The output weights are electrical and linear regression is performed electronically to extract symbols [24]. In our scheme, we propose to use optical weights for input, reservoir and output layers, that can further enhance the bandwidth. We utilize the nonlinear effects in SESAM (nonlinear amplitude change) or HNLW (nonlinear phase change) to realize the nonlinear activation function. Refs. [23] and [24] use discrete photonic components in the reservoir which increases the latency due to the longer lengths ($\sim$ hundred meters) of these devices. In our scheme, we propose to use integrated photonic platform with much shorter lengths ($<$1 mm) leading to reduced latency and real time equalization of the impairments. Sozos et al. have proposed an interesting equalization scheme that makes use of the four wave mixing (FWM) in highly nonlinear fiber (HNLF) [25] and this scheme does not employ RC system for the equalization.

Freiberger et al. proposed to use an integrated optical readout for the complex valued photonic RC which allows to perform all operations in optical domain [26], [27]. Masaad et al. have proposed a nonlinear equalization scheme based on photonic reservoir computing for the fiber optic system with Kramers-Kronig receivers which requires direct detection [28]. Although the schemes of [26], [27], [28] used complex weights, they are quite different from the proposed coherent RC: i) [26], [27] used direct detection and hence, the target function has to be real. In contrast, the target function of the coherent RC system is complex. Real valued target function is not useful for the equalization in coherent communication system since the signal in I-channel and Q-channel has to be equalized separately. ii) The schemes of [26], [27], [28] do not employ nonlinear devices in the optical domain. Instead, they use the nonlinear effect (square law detection) associated with O/E conversion in direct detection. Again, such a scheme cannot be used for the equalization in coherent communication systems because the O/E conversion in a coherent receiver is linear. iii) Due to the direct detection in [26], [27], [28], the states of the reservoir cannot be observed directly whereas in the case of coherent RC inserted in the coherent communication system, coherent receiver has access to both amplitude and phase of the received signal and hence, the states of the reservoir can be observed. In any case, the primary objective of this paper is to apply the RC concept to mitigate the linear and nonlinear impairments of the coherent fiber optic communication systems. Currently, the equalization of these impairments is carried out in the digital domain which is unattractive due to high power consumption and latency. Coherent RC allows the optical equalization of the

fiber impairments in real time with low power consumption and large bandwidth.

The paper is organized as follows. In Section II, a theoretical background on RC is provided. Various architectures such as electronic and optoelectronic RC systems are reviewed. Section III introduces a novel coherent all optical RC system just before the coherent receiver in a fiber optic communication system. In Section IV, numerical simulations of the coherent fiber optic communication system with and without the coherent RC are carried out. These simulations demonstrate the effectiveness of RC in mitigating the impairments. The performances of RC based on SESAM and HNLW are compared in long haul/short haul fiber optic communication systems.

II. Theoretical Background

Fig. 1 shows the schematic of a traditional RC system. Let $\mathbf{u}(n) = [u_1(n), u_2(n), \ldots, u_K(n)]^T$ be the input vector of dimension K where n is the discrete time. First, we construct N linear combinations of the input variables $u_1(n), u_2(n), \ldots, u_K(n)$ in the form

$$J_i(n) = \sum_{j=1}^{K} W_{ij}^{in} u_j(n), \qquad i = 1, 2, \ldots, N, \tag{1}$$

where W_{ij}^{in} are input weights that are randomly chosen from a simple distribution such as the Gaussian distribution and N is the number of nodes in the reservoir. Eq. (1) may be rewritten in the matrix form as

$$\mathbf{J}(n) = \mathbf{W}^{in}\mathbf{u}(n), \tag{2}$$

where $\mathbf{J}(n) = [J_1(n), J_2(n), \ldots, J_N(n)]^T$ and $\mathbf{W}^{in}$ is the input weight matrix. Let $x_j(n)$ represent the state of node j at the discrete time n. Let $y_m(n), m = 1, 2, \ldots, M$ be the reservoir outputs at time n. The evolution of the states of the reservoir is given by [3]

$$x_j(n) = f\left(\sum_{i=1}^{N} W_{ij,res}^{res} x_i(n-1) + \sum_{m=1}^{M} W_{mj,out}^{res} y_m(n-1) + J_j(n) + W_{j,bias}^{res}\right), \; j = 1, 2, \ldots, N, \tag{3}$$

where $W_{ij,res}^{res}$ is the weight for the recurrent connection between nodes i and j, $W_{mj,out}^{res}$ is the weight for the connection between output node m and reservoir node j, and $W_{j,bias}^{res}$ is the bias vector. $f(.)$ represents the nonlinear activation function. Often a sigmoidal function such as $f(x) = tanh(x)$ is chosen as the nonlinear activation funtion. The matrices $W_{ij,res}^{res}, W_{mj,out}^{res}$, and $W_{j,bias}^{res}$ are chosen randomly from a Gaussian distribution. These matrices are random but they do not change as a function of the discrete time n. In this paper, we do not include the feedback from the output nodes ($W_{mj,out}^{res} = 0$) and set the bias to zero ($W_{j,bias}^{res} = 0$). In this case, (3) becomes [3], [29]

$$x_j(n) = f\left(\sum_{i=1}^{N} W_{ij,res}^{res} x_i(n-1) + J_j(n)\right), \; j = 1, 2, \ldots, N \tag{4}$$

The reservoir has memory of past inputs. $x_j(n)$ are known as the internal states of the reservoir. The internal states of the reservoir are mapped to the reservoir outputs using

$$y_m(n) = \sum_{l=1}^{N} W_{ml}^{out} x_l(n), \qquad m = 1, 2, \ldots, M, \tag{5}$$

where W_{ml}^{out} are the output weights connecting the reservoir node l to the output m. Let the desired outputs be $\hat{y}_m(n), \quad m = 1, 2, \ldots M$. The output of the reservoir $y_m(n)$ should be as close as possible to the target function $\hat{y}_m(n)$. During the training period, the output weights W_{ml}^{out} are determined by minimizing the mean square difference between the output $y_m(n)$ and the target function $\hat{y}_m(n)$,

$$\begin{aligned} E_{ms}(\mathbf{W}^{out}) &= \frac{1}{LM}\sum_{m=1}^{M}\sum_{n=1}^{L}[\hat{y}_m(n) - y_m(n)]^2, \\ &= \frac{1}{LM}\sum_{m=1}^{M}\sum_{n=1}^{L}\left[\hat{y}_m(n) - \sum_{l=1}^{N} W_{ml}^{out} x_l(n)\right]^2, \end{aligned} \tag{6}$$

where L is the number of training samples and $\mathbf{W}^{out}$ is the output weight matrix with elements W_{ml}^{out}. The output weights W_{ml}^{out} can be found by setting the gradient to zero,

$$\frac{\partial E_{ms}}{\partial W_{kj}^{out}} = \frac{2}{LM}\sum_{n=1}^{L}\left[\hat{y}_k(n) - \sum_{l=1}^{N} W_{kl}^{out} x_l(n)\right] x_j(n) = 0, \quad j = 1, 2, \ldots N, \text{and } k = 1, 2, \ldots M. \tag{7}$$

Solving for W_{ml}^{out}, we find [30],

$$\mathbf{W}^{out} = \mathbf{x}^{\dagger}\hat{\mathbf{y}}, \tag{8}$$

where $\hat{\mathbf{y}}$ is a $L \times M$ matrix of the desired outputs,

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{y}_1(1) & \hat{y}_2(1) & \ldots & \hat{y}_M(1) \\ \hat{y}_1(2) & \hat{y}_2(2) & \ldots & \hat{y}_M(2) \\ . & . & \ldots & . \\ . & . & \ldots & . \\ . & . & \ldots & . \\ \hat{y}_1(L) & \hat{y}_2(L) & \ldots & \hat{y}_M(L) \end{bmatrix} \tag{9}$$

$\mathbf{x}$ is a $L \times N$ matrix of the internal states of the reservoir,

$$\mathbf{x} = \begin{bmatrix} x_1(1) & x_2(1) & \ldots & x_N(1) \\ x_1(2) & x_2(2) & \ldots & x_N(2) \\ . & . & \ldots & . \\ . & . & \ldots & . \\ . & . & \ldots & . \\ x_1(L) & x_2(L) & \ldots & x_N(L) \end{bmatrix} \tag{10}$$

and $\mathbf{x}^{\dagger}$ is the Moore-Penrose pseudo inverse of $\mathbf{x}$ given by

$$\mathbf{x}^{\dagger} = (\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T. \tag{11}$$

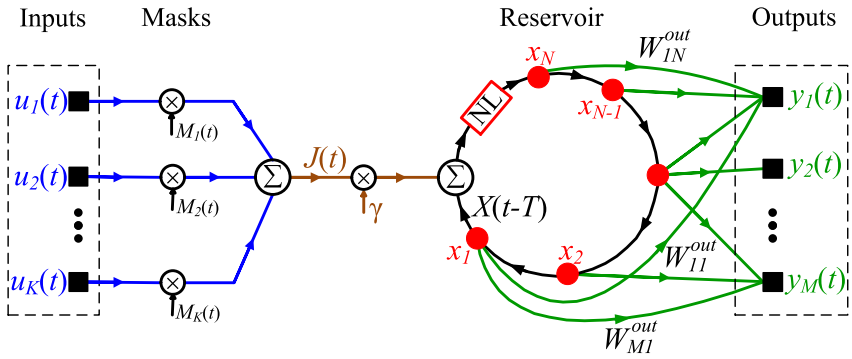


Fig. 2. Schematic of an RC system with a nonlinear node and delayed response. NL = nonlinear device.

To control overfitting, it may be necessary to add a regularization term to E_{ms} and the total error to be minimized takes the form

$$E_{tot}(\mathbf{W}^{out}) = E_{ms}(\mathbf{W}^{out}) + \lambda E_r(\mathbf{W}^{out}), \tag{12}$$

where λ is the regularization coefficient. One of the simplest regularizers is the sum-of-squares of the weight vector elements

$$E_r(\mathbf{W}^{out}) = \sum_{l=1}^{N}\sum_{m=1}^{M}(W_{lm}^{out})^2. \tag{13}$$

By setting the gradient of E_{tot} to zero, we find [30]

$$\mathbf{W}^{out} = (\lambda I + \mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\hat{\mathbf{y}}. \tag{14}$$

If a large number of training data is needed, the minimization of the RC training problem (6) can be more effectively done using any well-defined iterative gradient-based optimization method, e.g., the adaptive moment estimation (Adam) algorithm [31]. The Adam algorithm is one of the most effective stochastic-gradient-based optimization algorithms that is widely used to solve training problems of neural networks and deep neural networks. It is highly efficient in solving optimization problems with a large number of data and parameters, and with a little memory requirement.

A. RC With Nonlinear Node and Delayed Feedback

Appeltant et al. demonstrated the RC with a nonlinear node and delayed feedback [3]. Fig. 2 shows the block diagram of this scheme. Let the time required by the signal to traverse one loop be T. This delay is divided into N equal intervals of length, $\theta = T/N$. The N equidistant points in the loop can be interpreted as virtual nodes which play the role of nodes of the traditional reservoir. At time t_0, $\mathbf{u}(t_0) = [u_1(t_0), u_2(t_0), \ldots, u_K(t_0)]^T$ be the input vector. Input weight matrix $\mathbf{W}^{in}$ can be realized as follows. Consider a mask function

$$M_j(t) = W_{ij}^{in}, \quad \text{for } (i-1)\theta \le t \le i\theta, \quad i = 1, 2, \ldots N, \quad j = 1, 2, \ldots K, \tag{15}$$

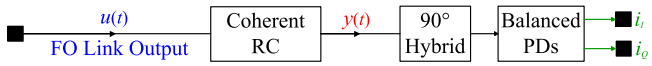


Fig. 3. Block diagram of a coherent receiver with coherent all optical RC.

and $M_j(t)$ is a periodic function with period T, $M_j(t+T) = M_j(t)$. Now, (1) maybe rewritten as

$$J(t) = \sum_{j=1}^{K} M_j(t)u_j(t), \tag{16}$$

where $J(t)$ plays the role of $J_i(n)$ with $t = (n-1)T + i\theta$, $i = 1, \ldots, N$. The internal states of the reservoir are represented by

$$X(t) \equiv x_j(n) = X((n-1)T + j\theta), \quad n = 1, 2, \ldots. \tag{17}$$

Here, $x_j(n)$ maybe interpreted as the internal state of the virtual node j at the discrete time n. The reservoir consists of a nonlinear device whose input is the sum of $\gamma J(t)$ and the delayed state $X(t-T)$ (see Fig. 2), and hence, its output is $h(\gamma J(t) + X(t-T))$ where γ is a suitably chosen factor known as input gain and $h(.)$ depends on the type of the nonlinear device used in the loop. The evolution of the internal state is typically governed by a first order differential equation of the form,

$$\frac{dX}{dt} = F(X(t), X(t-T) + \gamma J(t)). \tag{18}$$

On solving (18), internal states $X(t)$ can be found and using (17), internal states of the virtual nodes can be determined. The output vector is calculated by the weighted sum of the internal states (5). The training to determine W_{lm}^{out} is the same as before.

III. Coherent All Optical RC

Paquot et al. demonstrated an optoelectronic RC in which the delay loop is partly optical and partly electrical [4]. The optoelectronic RC [4] is not suitable for the equalization of impairments of coherent fiber optic networks. In optoelectronic RC, the inputs are in electrical domain. However, to take the full advantage of the fast optical signal processing, it is desirable to introduce the RC after the fiber optic (FO) link and just before the coherent receiver (see Fig. 3) and in this case, the inputs to the RC would be in the optical domain. So, the coherent RC may

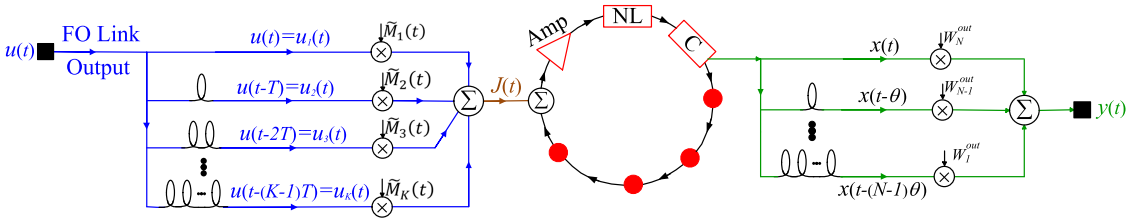


Fig. 4. Schematic of a coherent all optical RC system. Amp = amplifier, NL = nonlinear device, and C = coupler.

be treated as a part of the coherent receiver. The processing speed of the optoelectronic RC is limited by the long loop time T (∼ns) leading to a processing bandwidth of a few hundreds of MHz. However, to mitigate the impairments of FO system operating at 28 GBaud in real time, the processing speed of the RC must be of the order of a few tens of GHz which requires a loop time T in the order of a few picoseconds. Therefore, in this paper, we explore the possibility of using a coherent all-optical RC for the equalization of linear and nonlinear impairments of the FO system. Although the scheme of [26] introduces complex valued RC with high bandwidth and low power consumption, it is not suitable for the equalization of the impairments of a coherent fiber optic communication systems since the direct detection photodetectors are not compatible with coherent detection. The coherent all-optical RC combines the fast signal processing (such as matrix multiplications) in optical domain and short loop time of the delay loop implemented on photonic integrated circuits (PIC). It is believed that the short loop time of the delay loop implemented on PIC is a drawback of the photonic RC [29]. However, for equalization of impairments of the FO system operating at high data rates, delay loop on the PIC is an ideal candidate. Although shorter loop time leads to shorter memory, the memory can be enhanced by increasing the number of parallel inputs to the reservoir (see Fig. 4).

Fig. 3 shows the block diagram a fiber optic system with a coherent RC placed just before the coherent receiver. Fig. 4 shows the block diagram of the coherent all optical RC system. So far, we assumed that the inputs and outputs to be real. Now, consider the complex valued inputs and outputs. To equalize the inter symbol interference (ISI) caused by fiber dispersion and nonlinear effects, weighted sum of the delayed versions of the complex optical signal is fed to the reservoir. Let the input vector of the reservoir be $\mathbf{u}(t) = [u_1, u_2, \ldots u_K]^T$, where $u_j(t) = u(t-(j-1)T)$, and $\tilde{\mathbf{M}}(t) = [\tilde{M}_1, \tilde{M}_2, \ldots \tilde{M}_K]$. As before, mask functions are chosen from the Gaussian distribution. The difference between $\tilde{M}_j(t)$ and $M_j(t)$ of Section II-A is that $M_j(t)$ is real whereas $\tilde{M}_j(t)$ is complex. Now (16) can be rewritten as

$$J(t) = \tilde{\mathbf{M}}\mathbf{u}. \tag{19}$$

The vector multiplication of (19) can be realized on PIC [21]. The matrix multiplications on PIC can be done at the speed of light with a bandwidth exceeding 100 GHz [21], [32]. The delay lines in Fig. 4 can be realized using spirals in each waveguide and complex weights can be adjusted using the current controlled heaters [22]. Alternatively, voltage on the phase-shifters of a Mach-Zehnder interferometer (MZI) can be adjusted to obtain the desired complex weight [21]. Uncertainties associated with complex weights can be computed by uncertainty propagation [21]. The required delay T in Fig. 4 is of the order of picoseconds. If the delay is much less than the symbol period, it would have a negligible impact. However, if the delay is comparable to or larger than the symbol period, it can cause a significant penalty. The signal in the reservoir passes through an amplifier and a nonlinear device which corresponds to the nonlinear function $f()$ of (4). The input of the nonlinear device is

$$x_{in}(t) = \beta x(t-T') + \gamma J(t), \tag{20}$$

where T' is the round trip time chosen to be slightly different from the input time T so as to desynchronize the reservoir [4]. We chose $T' = T + \theta$.

TABLE I
PARAMETERS OF RC, SESAM AND HNLW

Nodes, N	8
Input time, T	35.71 ps
Round trip time, T'	40.17 ps
γ	0.01
β	0.05
k_1	2.3
k_0	0.5
P_0	0.09
α_1	2.3
α_0	0.009

We explore two types of nonlinear devices - (i) semiconductor saturable absorber mirror (SESAM), and (ii) highly nonlinear waveguide (HNLW). The nonlinear response of the SESAM is modeled as [33]

$$x(t) = Rx_{in}(t), \tag{21}$$

where R is the reflectivity,

$$R = \exp\left[-k_1 - \frac{k_0}{1+|x_{in}|^2/P_0}\right], \tag{22}$$

where k_1, k_0 and P_0 are the parameters of the SESAM which can be adjusted to approximate the experimental nonlinear response of SESAM of [10] or that of [21].

In the case of the HNLW, the nonlinear response is due to the Kerr effect. It is modeled as [34]

$$x(t) = \exp[-\alpha_1 + i\alpha_0|x_{in}|^2]x_{in}(t), \tag{23}$$

where α_1 and α_0 are the parameters of the HNLW. Table I shows the normalized parameters of SESAM and HNLW used in the numerical simulations. The normalization is done such that the maximum of $u(t)$ is unity. The difference between the SESAM and the HNLW is that the SESAM modifies the amplitude of the reservoir states while the HNLW modifies the phase of the states. Both SESAM and HNLW have dimensions less than 1 mm and hence, the latency associated with these devices is less than 10 ps. In contrast, the nonlinear activation functions associated with highly nonlinear fiber (HNLF) have larger dimensions ($\sim$100 m) and the latency is in the order of microseconds. Both SESAM and HNLW require relatively larger signal power ($\sim$ a few mW) to operate in the nonlinear regime.

The reservoir has got only one output $y(n)$ given by

$$y(n) = \sum_{l=1}^{N} W_l^{out} x_l(n). \tag{24}$$

The above equation may be written in the matrix form as

$$\mathbf{y} = \mathbf{x}\mathbf{W}^{out}, \tag{25}$$

where

$$\mathbf{y} = \begin{bmatrix} y(1) \\ y(2) \\ . \\ . \\ . \\ y(L) \end{bmatrix} \tag{26}$$

$$\mathbf{W}^{out} = \begin{bmatrix} W_1^{out} \\ W_2^{out} \\ . \\ . \\ . \\ W_N^{out} \end{bmatrix} \tag{27}$$

and $\mathbf{x}$ is given by (10).

To realize (24), a fraction of the signal in the reservoir is taken out using the coupler C (see Fig. 4). This signal is delayed by $l\theta$, $l = 1, 2, \ldots N-1$, multiplied by the corresponding weight W_l^{out}, and then summed up. Note that the output weights W_l^{out} are complex valued. During the training period, W_l^{out} can be determined by minimizing the mean absolute square difference between $y(n)$ and the desired output $\hat{y}(n)$,

$$\begin{aligned} E_{ms}(\mathbf{W}^{out}) &= \frac{1}{L}\sum_{n=1}^{L} |\hat{y}(n) - y(n)|^2 \\ &= \frac{1}{L}\sum_{n=1}^{L} \left| \hat{y}(n) - \sum_{l=1}^{N} W_l^{out} x_l(n) \right|^2. \end{aligned} \tag{28}$$

The weights W_l^{out} can be obtained by minimizing E_{ms}

$$\frac{\partial E_{ms}}{\partial W_j^{out}} = \frac{1}{L}\sum_{n=1}^{L} \left[\hat{y}^*(n) - \sum_{l=1}^{N} (W_l^{out})^* x_l^*(n) \right] x_j(n) = 0. \tag{29}$$

Solving (29), we find [26]

$$\mathbf{W}^{out} = (\mathbf{x}^H\mathbf{x})^{-1}\mathbf{x}^H\hat{\mathbf{y}}, \tag{30}$$

where the superscript H denotes the Hermitian transpose, $\hat{\mathbf{y}} = [\hat{y}(1), \hat{y}(2), \ldots \hat{y}(L)]^T$, and $\mathbf{x}$ is given by (10). The main difference between (11) and (30) is that the real transpose is replaced by the Hermitian transpose. As for regularization, since W_l^{out} is complex, we replace (13) by

$$E_r(\mathbf{W}^{out}) = \sum_{l=1}^{N} |W_l^{out}|^2. \tag{31}$$

The output weight vector $\mathbf{W}^{out}$ that minimizes E_{tot} is [26]

$$\mathbf{W}^{out} = (\lambda I + \mathbf{x}^H\mathbf{x})^{-1}\mathbf{x}^H\hat{\mathbf{y}}. \tag{32}$$

We note that the complex valued RC is more suitable for the equalization of impairments in the fiber optic systems since the optical field envelope is inherently complex.

The complex output of the coherent receiver is related to the reservoir output $y(t)$ by

$$\tilde{i}(t) \equiv i_I(t) + j i_Q(t) = R_c y(t) + n(t), \tag{33}$$

where $i_I(t)$ and $i_Q(t)$ are the in-phase and quadrature components of the current (see Fig. 3), R_c is a constant that depends on the parameters of the coherent detector such as responsivity, and $n(t)$ is the noise added by the coherent receiver. An important consequence of coherent detection is that the states of the reservoir can be estimated. During the training period, by choosing

$$\mathbf{W}^{out} = [1, 0 \ldots, 0]^T, \tag{34}$$

from (25), we find that the the first column of the state matrix can be estimated. Similarly, by choosing $\mathbf{W}^{out} = [0, 1 \ldots, 0]^T$, the second column of the state matrix can be estimated, and so on. In contrast, in the case of direct detection receiver, the reservoir states cannot be directly observed [26] since the receiver does not have access to the phase of the received signal which makes the training complicated.

Fiber optic channel is relatively stable and the channel may be considered as nearly static over a few milliseconds which is much longer than the symbol period ($\sim$ ps). Hence, the required frequency of training is much smaller than the baud rate. The values of the optimal output weights can be calculated in the digital domain and can be transferred to the optical domain as done in [21], [22]. As for the input weights, they can be drawn from a Gaussiand distribution in the digital domain and can be similarly transferred to the optical domain. As pointed in [21], phase encoding for complex weight implementation is likely to be the dominant source of error in the experimental realization of the proposed RC system.

IV. Equalization of Impairments of Coherent Fiber Optic System

In this Section, we explore the possibility of mitigating dispersive and nonlinear impairments in a coherent fiber optic system. There are two types of nonlinear distortions. (i) The transfer function of the MZM is not linear leading to distortion of the

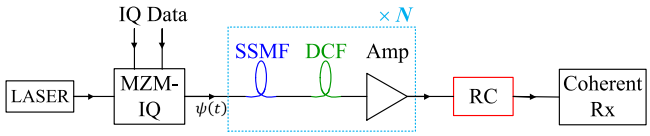


Fig. 5. Block diagram of a long haul dispersion managed fiber optic system. MZM = Mach-Zehnder modulator, IQ = inphase-quadrature, SSMF = standard single mode fiber, DCF = dispersion compensation fiber, Amp = Erbium-doped fiber amplifier, and RC = complex valued all optical reservoir computing.

signal pulses launched to the fiber. (ii) Due to the Kerr effect in fibers, the optical signal undergoes self-phase modulation (SPM) [34], intra-channel cross-phase modulation (IXPM) [35], and intra-channel four wave mixing (IFWM) [36], [37], [38], [39]. In Section IV-A, it is shown that the RC can compensate for these nonlinear effects. The pulses broaden due to fiber dispersion and interfere with neighboring symbols leading to performance degradations. Section IV-B shows that the RC can compensate for the signal distortion due to dispersion.

A. Equalization of Nonlinear Impairments in Long Haul Systems

Consider a long haul fiber optic system shown in Fig. 5. The inphase-quadrature (IQ)- Mach-Zehnder modulator (MZM) is driven by the electrical IQ data and the output of IQ-MZM passes through a dispersion managed (DM) fiber optic link. The DM link consists of a standard single mode fiber (SSMF) followed by a dispersion compensation fiber (DCF) so that the dispersion is fully compensated for at each span. The complex-valued RC is introduced at the receiver just before the coherent detection. In this system, the RC compensates for both the nonlinear transfer function of MZM-IQ and the Kerr nonlinearity of TF and DCF. The outputs of the I and Q branches of the MZM-IQ are given by [40]

$$\psi_I(t) = \cos\left[\frac{(m_I(t) - V_{bias}/2)\pi}{V_\pi}\right], \tag{35}$$

$$\psi_Q(t) = \cos\left[\frac{(m_Q(t) - V_{bias}/2)\pi}{V_\pi}\right], \tag{36}$$

where $m_I(t)$ and $m_Q(t)$ are the electrical in-phase and quadrature data, V_π is the switching voltage and V_{bias} is the bias voltage. The complex field envelope of the optical field output of the MZM-IQ is

$$\psi = \psi_I(t) + i\psi_Q(t). \tag{37}$$

Numerical simulations of a single channel and single polarization fiber optic communication system with complex-valued RC placed just before the coherent receiver is carried out with the following parameters: Data rate = 28 GBaud, and modulation format = quadrature phase shift keying (QPSK). The well known split-step Fourier scheme [34] is used to simulate the nonlinear propagation in fibers. Total transmission distance = 3200 km. The message signals, $m_I(t)$ and $m_Q(t)$ are pseudo-random bit sequences of length $2^{16} - 1$. The raised cosine pulses with a roll-off factor of 0.1 are used for both $m_I(t)$ and $m_Q(t)$. The coherent receiver adds shot noise and thermal noise. The

TABLE II
SIGNAL AND SYSTEM PARAMETERS

IQ-MZM, V_{bias}	4 V
IQ-MZM, V_π	4 V
Tx laser linewidth	20 kHz
No. of symbols	409600
SSMF, D	17 ps/nm/km
SSMF, Loss	0.2 dB/km
SSMF, Length	80 km
SSMF, NL Coeff. γ_{SSMF}	1.1 $W^{-1}km^{-1}$
DCF, D	-100 ps/nm/km
DCF, Loss	0.45 dB/km
DCF, NL Coeff. γ_{DCF}	4.4 $W^{-1}km^{-1}$
EDFA, NF	4.77 dB
Transmission Reach	3200 km
LO power	10 mW
LO linewidth	20 kHz
Load Resistance	1 kΩ
Temperature	290 Kelvin

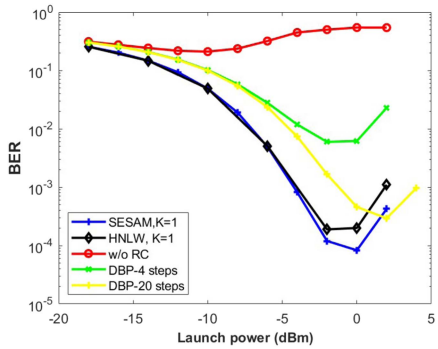


Fig. 6. Plot of the fiber launch power versus BER in a long haul fiber optic system without RC, with RC and with DBP. DBP-X steps corresponds to the DBP with X steps per fiber span.

details of the implementation of the coherent receiver with balanced photodetectors can be found in [40]. The phase noise due to transmitter laser and LO are compensated for using the Viterbi-Viterbi algorithm [41]. The parameters of IQ-MZM, TF, DCF and coherent receiver are listed in Table II. During the training mode, 4096 symbols were transmitted. For a specific choice of the random matrix $\tilde{\mathbf{M}}$, the optimum output weights are calculated and the BER is measured. This procedure is repeated 20 times by changing the random matrix $\tilde{\mathbf{M}}$ and the corresponding optimum output weights are calculated. The random matrix and the corresponding output weights that provides the lowest BER is retained and during the data transmission mode, this random matrix and output weights are used. Fig. 6 shows the BER as a function of the fiber launch power without RC, with RC based on SESAM, with RC based on HNLW, and with digital back propagation (DBP). First consider the system without RC. When the launch power is very low (< -14 dBm), the BER is high because the noise power due to inline amplifiers is large (i.e. signal-to-noise ratio (SNR) is low). As the launch power increases, the BER improves slightly and then it becomes higher at higher launch powers (> -6 dBm) because the fiber nonlinear effects kick in leading to nonlinear impairments. Next, consider the fiber optic system with RC based on SESAM ($K = 1$). It can be seen that the BER is much lower than that of the system

without RC even at high launch powers (> -3 dBm) demonstrating that the RC is quite effective in mitigating the nonlinear impairments. The RC based on HNLW also provides significant performance benefit as compared to the system without RC. However, the RC based on SESAM outperforms that based on HNLW. The lowest BER for the case of the system without RC is 0.209 occuring at a launch power of -10 dBm which decreases to 1.89×10^{-4} at -2 dBm using the RC based on HNLW. The BER can be further reduced to 8.29×10^{-5} at 0 dBm using the RC based SESAM. Thus, the nonlinear tolerance is increased by 8 dB and 10 dB using HNLW and SESAM, respectively, as compared to the system without RC. In Fig. 6, the performance of the fiber optic system with RC is also compared with the system with the DBP. In the case of DBP, virtual fibers are realized in the digital domain using the split step Fourier scheme (SSFS) [42], [43]. When we use the DBP with four steps per fiber span, the minimum BER is 5.9×10^{-3}, which is reduced to 2.9×10^{-4} using the DBP with 20 steps per fiber span. In Fig. 6, we see that the system with RC (SESAM or HNLW) is slightly better than the DBP with 20 steps per span. The number of fast Fourier transform (FFT) pairs required for the DBP is directly proportional to the number of steps per span and hence, the computational cost of the DBP increases with the numbers of steps per span. For example, for a 3200 km dispersion managed system, if we use 20 steps per span, the number of FFT pairs would be 1600! Therefore, it would be very hard to realize the DBP with a large number of steps per span on the DSP in real time and most of the DBP realizations are limited to off-line signal processing. In contrast, the coherent RC can be realized in real time. In addition, because of the extensive computations required for the DBP, the power consumption of the DSP and the latency scale directly with the number of steps per span.

So far we assumed that $K = 1$ which means we do not include the delayed versions of the fiber optic link output as inputs to the RC system (see Fig. 4). As the optical signal propagates through the dispersion managed fiber, the pulses broaden due to dispersion and the signal pulse in the current symbol slot interacts nonlinearly with the neighboring symbols. To counteract this effect, the RC should have a long nonlinear memory. From Fig. 6, we infer that the RC based on SESAM has a longer memory as compared to that based on HNLW. The nonlinear memory can be enhanced by introducing the delayed versions of the fiber optic link outputs as inputs to the RC system. Fig. 7 compares the BER of the RC based on SESAM and HNLW for $K = 1$ and $K = 3$. As can be seen, the BER can be improved by increasing K for both RC systems. Fig. 8(a) and (b) compare the constellation diagrams of the fiber optic system without RC and with RC based on SESAM when the fiber nonlinear effects are very strong (launch power = 5 dBm). The constellation diagram of the RC based on HNLW is quite similar to Fig. 8(b).

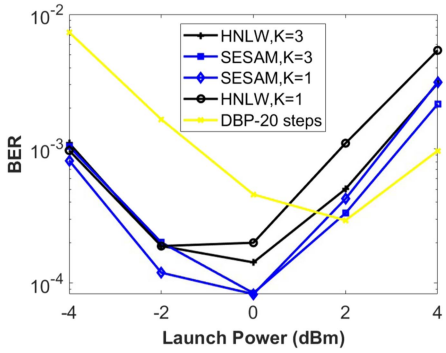
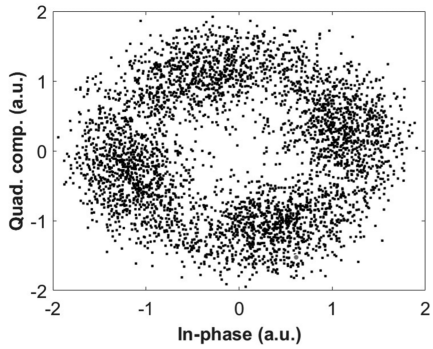
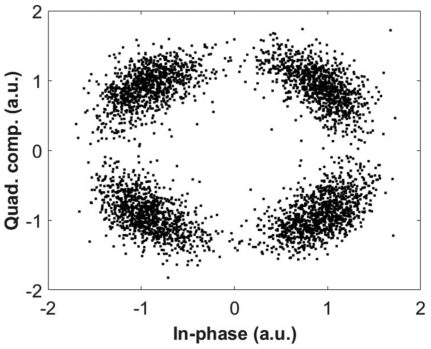


Fig. 7. Comparison of RC schemes based on SESAM and HNLW.



(a)



(b)

Fig. 8. Constellation diagrams for a long haul fiber optic system (a) without the RC and (b) with the SESAM. Fiber launch power = 5 dBm.

B. Equalization of Dispersive Impairments in Short Haul Systems

Consider a short haul fiber optic system such as that used for data center networks. Fig. 9 shows the block diagram of such a system with the complex valued RC placed just before the coherent receiver. The output of the MZM-IQ passes through an SSMF of length 80 km. The rest of the signal and system parameters are the same as that of Tables I and II. In this system, the RC compensates for the MZM nonlinearity, the fiber dispersion, and the fiber nonlinearity. Since the fiber length is short, the fiber nonlinear effect is not significant and the distortion is mainly due to fiber dispersion. Fig. 10 shows the BER as a function of the fiber launch power. In this FO system, a signal pulse broadens and overlaps linearly with the neighboring signal pulses. In this case, we need an RC system that has a longer linear memory. First, consider the case $K = 1$. From Fig. 10, we see that the RC based on HNLW outperforms that

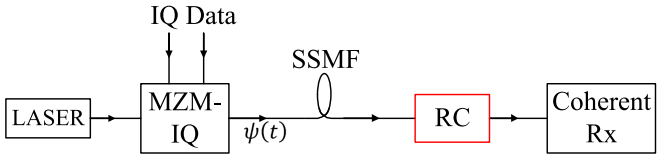


Fig. 9. Block diagram of a short haul fiber optic system. MZM = Mach-Zehnder modulator, IQ = inphase-quadrature, SSMF = standard single mode fiber, and RC = complex valued all optical reservoir computing.

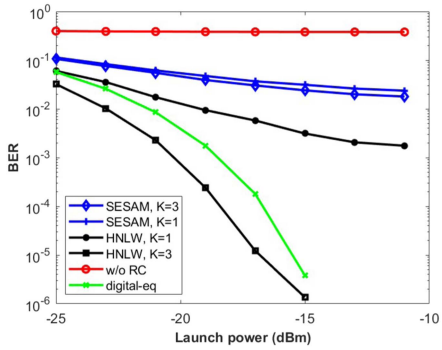
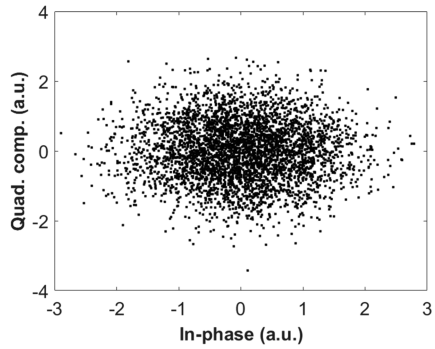
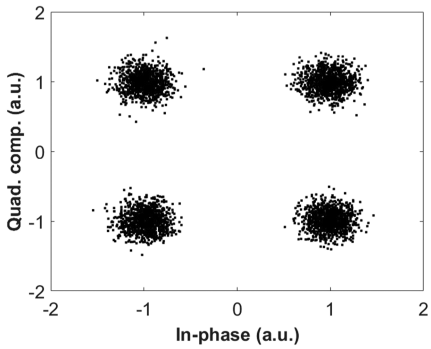


Fig. 10. Plot of the fiber launch power versus BER in a short haul fiber optic system with and without RC.



(a)



(b)

Fig. 11. Constellation diagrams for a short haul fiber optic system (a) without RC and (b) with the RC based on HNLW. Received power = −15 dBm. $K = 3$.

based on SESAM which implies that the RC based on HNLW has a longer linear memory. However, neither of them can fully compensate for fiber dispersion. When $K = 3$, the performance of the RC based on SESAM improves slightly whereas that of the RC based on HNLW improves drastically. For $K = 3$ and with HNLW, the dispersion can be fully compensated for. If the launch power exceeds -15 dBm, the BER becomes smaller than 10^{-6}. Figs. 11(a) and (b) show the constellation diagrams without RC and with the RC based on HNLW, respectively. As can be seen, the RC based on HNLW can effectively mitigate the dispersive impairments. In Fig. 10, we compare the performance of the system based on RC with that uses a digital equalizer. Inverse transfer function of the transmission fiber is realized in the digital domain and applied to the received signal in the frequency domain [40]. As can be seen in Fig. 10, the performance of the system with HNLW ($K = 3$) is slightly better than that with the digital equalizer. The likely reason is that the system with HNLW compensates for both linear and nonlinear impairments (although nonlinear impairments are small in this system) whereas the digital equalizer compensates only for the linear (chromatic dispersion) impairment.

V. Conclusion

We have shown that a coherent all optical RC placed just before the coherent receiver is a promising technique for mitigating linear and nonlinear impairments of a coherent fiber optic communication system. The previous implementations of the RC for the mitigation of impairments in coherent fiber optic system have either used the electrical signal after the optical receiver as the input of the reservoir or used input/output weights in the electrical domain and hence, the processing speed is limited by electronic bottleneck. We have explored the performance of an RC system with either SESAM or HNLW as the nonlinear device in the reservoir. The SESAM provides a nonlinear change in amplitude whereas the HNLW provdies the nonlinear phase change. Both SESAM and HNLW based RC systems provide significant performance improvement by mitigating the dispersive and nonlinear impairments of a fiber optic system as compared to the system without RC. For a dispersion managed long haul fiber optic system, the dominant impairment is due to fiber nonlinear effects and for this system, the SESAM based RC is found to outperform the HNLW based RC since the former has a larger nonlinear memory. For a short haul fiber optic system, nonlinear effects are negligible and the dominant impairment is due to fiber dispersion. For this system, HNLW based RC has a far superior performance compared to the SESAM since the former has a larger linear memory.

The maximum dimension of the proposed RC system is less than 1 mm and hence, the latency introduced by the RC system is less than 10 ps whereas the all-optical RC systems that employ discrete photonic components such as highly nonlinear fibers have the much longer latency ($\sim$ microseconds). In addition, in the proposed RC design, the loop time of the reservoir is of the order of the symbol period enabling the real time equalization of the impairments in high data rate coherent optical communication systems. In order to realize the proposed design experimentally, (i) an optical delay and multiply circuit is required which has been experimentally demonstarated in [22], and (ii) the reservoir loop should have a nonlinear device such as

SESAM or HNLW to serve as the nonlinear activation function. These devices are already available. The challenge is to integrate all of these integrated photonic devices and to ensure that the specifications on the delay and complex weights are met. The phase encoding for the complex weight implementation is likely to be the dominant source of error [21] in the experimental realization of the proposed RC system.

REFERENCES

[1] H. Jaeger, "The "echo state" approach to analysing and training recurrent neural networks-with an Erratum note," Bonn, Germany: *German Nat. Res. Center Inf. Technol.*, GMD Tech. Rep. no. 148, vol. 148, no. 34, 2001, Art. no. 13.

[2] 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, 2002.

[3] L. Appeltant et al., "Information processing using a single dynamical node as complex system," *Nature Commun.*, vol. 2, no. 1, 2011, Art. no. 468.

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

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

[6] W. Du et al., "An optoelectronic reservoir computing for temporal information processing," *IEEE Electron Device Lett.*, vol. 43, no. 3, pp. 406–409, Mar. 2022.

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

[8] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, "Parallel photonic information processing at gigabyte per second data rates using transient states," *Nature Commun.*, vol. 4, no. 1, 2013, Art. no. 1364.

[9] C. Mesaritakis, V. Papataxiarhis, and D. Syvridis, "Micro ring resonators as building blocks for an all-optical high-speed reservoir-computing bit-pattern-recognition system," *J. Opt. Soc. Amer. B*, vol. 30, no. 11, pp. 3048–3055, 2013.

[10] A. Dejonckheere et al., "All-optical reservoir computer based on saturation of absorption," *Opt. Exp.*, vol. 22, no. 9, pp. 10868–10881, 2014.

[11] M. Nakajima, M. Inubushi, T. Goh, and T. Hashimoto, "Coherently driven ultrafast complex-valued photonic reservoir computing," in *Proc. Conf. Lasers Electro-Opt.*, 2018, Art. no. SM1C–4.

[12] K. Vandoorne et al., "Experimental demonstration of reservoir computing on a silicon photonics chip," *Nature Commun.*, vol. 5, no. 1, 2014, Art. no. 3541.

[13] Q. Vinckier et al., "High performance photonic reservoir computer based on a coherently driven passive cavity," *Optica*, vol. 2, no. 5, 2015, Art. no. 438.

[14] A. Argyris, J. Bueno, and I. Fischer, "Photonic machine learning implementation for signal recovery in optical communications," *Sci. Rep.*, vol. 8, no. 1, 2018, Art. no. 8487.

[15] A. Argyris, J. Bueno, and I. Fischer, "PAM-4 transmission at 1550 nm using photonic reservoir computing post-processing," *IEEE Access*, vol. 7, pp. 37017–37025, 2019.

[16] S. Li, S. Ohlendorf, and S. Pachnicke, "100 km 56 GBD PAM-4 transmission using photonic reservoir computing," in *Proc. IEEE 45th Eur. Conf. Opt. Commun.*, 2019, pp. 1–4.

[17] A. Katumba, X. Yin, J. Dambre, and P. Bienstman, "A neuromorphic silicon photonics nonlinear equalizer for optical communications with intensity modulation and direct detection," *J. Lightw. Technol.*, vol. 37, no. 10, pp. 2232–2239, May 2019.

[18] F. Da Ros, S. M. Ranzini, H. Bülow, and D. Zibar, "Reservoir-computing based equalization with optical pre-processing for short-reach optical transmission," *IEEE J. Sel. Topics Quantum Electron.*, vol. 26, no. 5, Sep./Oct. 2020, Art. no. 7701912.

[19] X. Zuo et al., "Integrated silicon photonic reservoir computing with PSO training algorithm for fiber communication channel equalization," *J. Lightw. Technol.*, vol. 41, no. 18, pp. 5841–5850, Sep. 2023.

[20] Y. Osadchuk, O. Jovanovic, D. Zibar, and F. Da Ros, "Multi-symbol reservoir computing-based equalization for PAM-4 IM/DD transmission," *IEEE Photon. Technol. Lett.*, vol. 36, no. 13, pp. 853–856, Jul. 2024.

[21] Y. Shen et al., "Deep learning with coherent nanophotonic circuits," *Nature Photon.*, vol. 11, no. 7, pp. 441–446, 2017.

[22] E. Staffoli, M. Mancinelli, P. Bettotti, and L. Pavesi, "Equalization of a 10 Gbps IMDD signal by a small silicon photonics time dealyed neural network," *Photon. Res.*, vol. 11, no. 5, 2023, Art. no. 878.

[23] M. Sorokina, S. Sergeyev, and S. Turitsyn, "Fiber echo state network analogue for high bandwidth dual quadrature signal processing," *Opt. Exp.*, vol. 27, no. 3, 2019, Art. no. 2387.

[24] M. Sorokina, "Dispersion managed fiber echo state network analogue with high (including THz) bandwidth," *J. Lightw. Technol.*, vol. 38, no. 12, pp. 3209–3213, Jun. 2020.

[25] K. Sozos, S. Deligiannidis, C. Mesaritakis, and A. Bogris, "Unconventional computing based on four wave mixing in highly nonlinear waveguides," *IEEE J. Quantum Electron.*, vol. 60, no. 4, Aug. 2024, Art. no. 9100106.

[26] M. Freiberger, A. Katumba, P. Beinstman, and J. Dambre, "Training passive photonic reservoirs with integrated optical readout," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, pp. 1943–1953, Jul. 2019.

[27] M. Freiberger, "Scaling up integrated photonics reservoirs towards low-power high-bandwidth computing," Ph.D. dissertation, Universiteit Ghent, Gent, Belgium, 2020.

[28] S. Masaad, E. Gooskens, S. Sackesyn, J. Dambre, and P. Bienstman, "Photonic reservoir computing for nonlinear equalization of 64 QAM signals with a Kramers-Kronig receiver," *Nanophotonics*, vol. 12, no. 5, 2023, Art. no. 925.

[29] G. Tanaka et al., "Recent advances in physical reservoir computing: A review," *Neural Netw.*, vol. 115, pp. 100–123, 2019.

[30] C. M. Bishop and N. M. Nasrabadi, *Pattern Recognition and Machine Learning*, vol. 4, no. 4. Berlin, Germany: Springer, 2006.

[31] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Representations*, Dec. 22, 2014.

[32] L. Vivien et al., "Zero-bias 40Gbit/s germanium waveguide photodetector on silicon," *Opt. Exp.*, vol. 20, no. 2, pp. 1096–1101, 2012.

[33] T. Hirooka, K. Tokuhira, M. Yoshida, and M. Nakazawa, "440 FS, 9.2 GHz regeneratively mode-locked erbium fiber laser with a combination of higher-order solitons and a SESAM saturable absorber," *Opt. Exp.*, vol. 24, no. 21, pp. 24255–24264, 2016.

[34] G. P. Agrawal, "Nonlinear fiber optics," in *Nonlinear Science at the Dawn of the 21st Century*. Berlin, Germany: Springer, 2000, pp. 195–211.

[35] S. Kumar, J. C. Mauro, S. Raghavan, and D. Q. Chowdhury, "Intrachannel nonlinear penalties in dispersion-managed transmission systems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 8, no. 3, pp. 626–631, May/Jun. 2002.

[36] A. Mecozzi, C. B. Clausen, and M. Shtaif, "Analysis of intrachannel nonlinear effects in highly dispersed optical pulse transmission," *IEEE Photon. Technol. Lett.*, vol. 12, no. 4, pp. 392–394, Apr. 2000.

[37] R.-J. Essiambre, G. Raybon, and B. Mikkelsen, "Pseudo-linear transmission of high-speed TDM signals: 40 and 160 Gb/s," in *Optical Fiber Telecommunications IV-B*. San Francisco, CA, USA: Academic, 2002, pp. 232–304.

[38] T. Hirooka and M. J. Ablowitz, "Analysis of timing and amplitude jitter due to intrachannel dispersion-managed pulse interactions," *IEEE Photon. Technol. Lett.*, vol. 14, no. 5, pp. 633–635, May 2002.

[39] D. Yang and S. Kumar, "Intra-channel four-wave mixing impairments in dispersion-managed coherent fiber-optic systems based on binary phase-shift keying," *J. Lightw. Technol.*, vol. 27, no. 14, pp. 2916–2923, Jul. 2009.

[40] S. Kumar and M. J. Deen, *Fiber Optic Communications: Fundamentals and Applications*. Hoboken, NJ, USA: Wiley, 2014.

[41] A. J. Viterbi and A. M. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmissions," *IEEE Trans. Inf. Theory*, vol. 29, no. 4, pp. 543–551, Jul. 1983.

[42] E. Ep and J. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *J. Lightw. Technol.*, vol. 26, no. 20, pp. 3416–3425, Oct. 2008.

[43] X. Li et al., "Electronic post-compensation of WDM transmission impairments using coherent detection and digital signal processing," *Opt. Exp.*, vol. 16, no. 2, pp. 880–888, 2008.