



Fiber echo state network analogue for high-bandwidth dual-quadrature signal processing

MARIIA SOROKINA, SERGEY SERGEYEV, AND SERGEI TURITSYN

Aston University, B4 7ET, United Kingdom

*m.sorokina@aston.ac.uk

Abstract: All-optical platforms for recurrent neural networks can offer higher computational speed and energy efficiency. To produce a major advance in comparison with currently available digital signal processing methods, the new system would need to have high bandwidth and operate both signal quadratures (power and phase). Here we propose a fiber echo state network analogue (FESNA) — the first optical technology that provides both high (beyond previous limits) bandwidth and dual-quadrature signal processing. We demonstrate applicability of the designed system for prediction tasks and for the mitigation of distortions in optical communication systems with multilevel dual-quadrature encoded signals.

© 2019 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Machine learning, in particular neural networks (NNs), have recently attracted renewed and fast-growing interest due to novel designs, advanced hardware implementations, and new applications. The surge of interest in recurrent neural networks (RNNs) is due to the broad spectrum of successful applications, which ranges from image analysis and classification, speech recognition, and language translation to more specific tasks like solving inverse imaging problems or signal processing in wireless and optical communications. In particular, in optical communications, RNNs have been used to compensate for nonlinear signal distortions [1–3] and nonlinearity mitigation for QPSK and 16-QAM modulation formats [4–6]. However, many applications are limited because the typical scale of RNNs requires a significant amount of time for training and processing, which represents a challenging computational burden. Various electronic and optical hardware designs have been proposed and developed recently to enable faster implementation of RNNs.

The first realistic silicon neuron was built in the 1990s [7] and there is now a silicon version of the eye's retina [8] and the Neurogrid network that simulates one million neurons [9]. Similarly, developed within the EC Human Brain Project (HBP) advanced silica-wafer-based SpiNNaker and BrainScaleS prototypes, also used external memory array. Also, in [10] silicon photonic weight banks were used to solve a differential system predicting 294-fold acceleration against a conventional computer. Only in semiconductor NN, the memory is achieved by optical delayed feedback [11]. Fiber-based NN was demonstrated recently with gigabyte per second speed [12], where memory was introduced by external electrical memory array - this imposed a limit on the design, as the learning could not be done in real time.

One of the promising types of RNN is Echo State Network (ESN) [13], a specific type of reservoir computing (RC), which relaxes training complexity. RC requires only output weights to be trained and can be realized by randomly connected nonlinear nodes, "neurons," while the training is achieved by collecting the signals at each state to compute optimum output weights, moreover, ESN does not have limitations on the pulse shape. Non-linear equalization based on RC has been proposed [14]. The "reservoir"-based approach makes it possible to demonstrate network comprising only *one nonlinear node and delay line*. All-optical implementation of

RC enables high-speed signal processing and can set the framework for a new generation of hardware [11, 15] for computing and future optical networks. One of the first demonstrations of an all-optical RC utilizing GHz bandwidth was based on a semiconductor laser [11], which was also used recently for amplitude modulated-signal processing [16]. Yet, even the most advanced realizations are fundamentally limited by the operation bandwidth of the order of a few GHz [11, 17] (due to the relaxation time of the media) and focus on signal power transformations, while fiber can unlock processing of tens of THz order bandwidth and enables dual-quadrature transformations.

Here we introduce a new approach to reservoir computing exploiting fiber nonlinearity, which makes it possible to receive nonlinear "neuron" functionality via signal-pump beating. Thus, memory is intrinsic to the process and is realized optically. We propose and demonstrate through numerical modeling a novel design of fiber echo state network analogue (FESNA). The advantages of the proposed setup include: (i) NN with intrinsic optical memory (does not require an external memory as in [12]) and (ii) for the first time possibility of optical RNN is shown for high-bandwidth (over 100 GHz compared to previous limits of a few GHz (e.g. [11])) and dual-quadrature signal processing. These features are highly important for increasing the speed and applicability of the optical neuromorphic technology. We demonstrate the feasibility of the proposed system using the examples of two practical applications dealing with predictions and distortion mitigation tasks. Our analysis reveals the device's strong potential for signal processing in high-speed optical communications. The presented all-optical RNN implementation is transparent and flexible and can find a wide variety of applications, ranging from photonics to medicine [18], in particular, RNN for self-tuning lasers [19, 20], smart fiber gratings [21], off-the-shelf smart fiber couplers [22], and so on.

2. Method

2.1. Reservoir computing concept

Reservoir computing is a type of recurrent network with randomly connected nonlinear nodes, where only the output weights are updated via training. This significantly simplifies the structure and training time. It has been shown that reservoir computing operation can be realized by a single nonlinear element with a single feedback delay loop [23].

Similar to neural networks, here one operates with nodes, which mimic neurons. The virtual neurons are obtained by sampling the input signal u_n . The samples are then combined with random input weights W^{in} . Consequently, each sample u_n is transformed into the weighted mix of the sequence $W^{in}u_n$, thus is performed the mapping of the signal into multidimensional space. To introduce memory into the system, the signal is also mixed with the time-delayed feedback. The signal from the feedback loop is the output of nonlinear element, which is also sampled with the same sampling rate - this sequence represents the state of the virtual network x_{n-1} , which is also mixed through a different weight matrix W . The feedback enables the memory in the network, where each state is affected by previous one. The received signal $Wx_{n-1} + W^{in}u_n$ represents the signals processed by the neuron n , coming from adjunct synapses with different weights (strength). The nonlinear response of the neuron is modeled by the nonlinear response function $y = f(x)$ (Fig. 1(b)). Thus, one receives nonlinear neurons processing input signals.

Meanwhile, the output of the feedback loop is also gradually collected at the receiver $X = [\dots x_n, x_{n+1}, \dots]$. Here, the weights of the matrices W^{in} and W were chosen randomly from the $[-0.5, 0.5]$ interval, thus, having uniform distribution around zero on the given interval. While, for processing the real part of the signal, one can use real-valued weight matrices, for processing complex signals the weight matrices were chosen similarly but in the complex domain. Once the weights are chosen, they remain fixed.

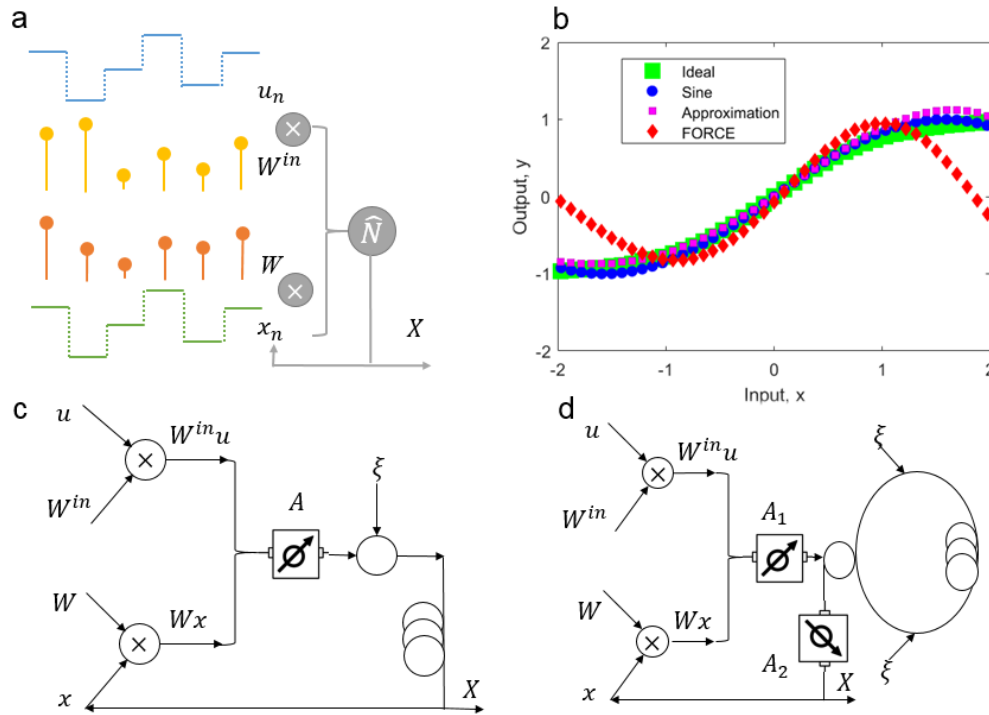


Fig. 1. The proposed reservoir setup. (a) Reservoir computing basics; (b) Transfer functions: ideal \tanh -function (green squares) and approximations by sin-function (blue circles), analytical approximation (violet squares) and FESNA-generated (red diamonds). (c) 1D-processing setup with parameters: 1W pump power ($\xi = 1$), nonlinear fiber parameters $|\xi|^2 \gamma L = 2.2\pi$, attenuation coefficient $A = -22.5\text{dB}$. A 3dB coupler is assumed. (d) 2D-processing setup with parameters: pump ($\xi = 1 + 1i$), nonlinear fiber parameters $|\xi|^2 / 2\gamma L = 2.2\pi$, attenuation coefficients ($A_1 = -18.5\text{dB}$, $A_2 = -6\text{dB}$).

Overall, the governing equation of the network:

$$x_n = f(Wx_{n-1} + W^{in}u_n) \quad (1)$$

where x_n is the N -dimensional reservoir state, f is a sigmoid function (usually the logistic sigmoid or the \tanh function), W is the $N \times N$ reservoir weight matrix, W^{in} is the $N \times K$ input weight matrix, u_n is the K -dimensional input signal. The training data u are mixed with the mask W^{in} and fed into the reservoir, where it is mixed with building up signal x mixed with a different mask W (Fig. 1(a)).

Here and throughout the paper, we focus on achieving nonlinear neural-resembling response - that is, sigmoid activation function - which is typically used in RNNs; therefore, we assume weights (W and W^{in} to be applied electronically, similar to [16] or [24–26]). Note, while previous configurations were limited fundamentally by a few orders of GHz [11] due to the relaxation time of the medium, fiber-based technology enables processing of THz bandwidth signals. Therefore, the proposed system is limited only by the electronic weight application, which currently enables the order of hundred GHz signal bandwidth processing. Further work will include modeling weights optically using physical effects in fiber. In the present study, however, we focus solely on realizing nonlinear neural functionality via fiber, thus unveiling a new regime in fiber systems that can be achieved under specific conditions described below and not accessible with traditional techniques.

The reservoir size N is a number of virtual nodes in the reservoir loop - optical feedback loop implemented via fiber. This can be achieved by spatial multiplexing or temporal multiplexing. One can use the temporal multiplexing approach of [11]. This causes the delay in the loop, as we feed each symbol of u_n for each x_n -vector with size of the reservoir. Overall, this realization results in N/N_s -delay per symbol, where N_s -is the signal sampling rate. This can be avoided by using spatial multiplexing; that is, a multi-core fiber where signals will be spread over N -core fiber. Such an approach will make it possible to avoid delay loop. Further for simplicity we use time delay reservoir with $N = 256$, which was found to be optimal number for the considered two tasks of prediction and distortion mitigation (classification).

At the final stage all the states of the reservoir are collected at matrix X . Then a simple linear regression or pseudo-inversion can be used to obtain the optimum weights W^{out} ; this can be done offline or online [27]. Thus, complex signal processing is reduced to linear operation, while the complexity is transferred to optical domain.

Once the output weights have been identified, the reservoir is tailored for a specific task and can be used in a straightforward manner. Here we will apply it in two examples: a standard prediction task (Mackey-Glass test) and for optical communication processing. In the first instance we use a single-quadrature 100 GHz-bandwidth signal and in the second case we consider multilevel dual-quadrature encoded signals with typical for optical communications channel bandwidth of 30 GHz. This will demonstrate the transparency of the technology to a different bandwidth range. We will also show configurations for one-dimensional (amplitude) and two-dimensional (both quadratures) signal processing (in Fig. 1, panels (c) and (d), respectively), while the underlying principle is the same, using Kerr nonlinearity to achieve nonlinear transformation and pump to select the required quadrature. Note, here we assume that weights and sampling are applied electronically (see [11] for more information), focusing on the optical realization of nonlinear functionality. Fiber dispersion can be used for mixing signals with weights optically in a temporal-multiplexing scenario and fiber core coupling can be used in the spatial-multiplexing case. However, here for the simplicity of presenting the concept, we consider all fibers to have insignificant dispersion; for example, operating near the zero dispersion point. In modern fibers, high nonlinearity and negligible dispersion coefficients can be achieved at the considered span lengths.

2.2. Setup for single-quadrature signal processing

The nonlinear sigmoid functionality (function, f) (Fig. 1(b)) can be approximated using the sine function. The latter can be achieved via Kerr nonlinearity in nonlinear fiber 1(c). By coupling the signal with the modulated pump and carefully optimizing fiber parameters, sine transformation can be achieved over the input signal.

For simplicity, let us assume a dispersionless nonlinear fiber, which will result in a strong power-dependent phase shift due to self-phase modulation (SPM) effect. Using dispersionless approximation, we can derive simple analytical expressions. After coupling of the weighted signal and reservoir output,

$$x_{n+1}^{(1)} = Wx_n + W^{in}u_{n+1} \quad (2)$$

the signal is attenuated (with attenuation coefficient A) to ensure the pump is much stronger than the signal, even for reasonable pump powers. Higher pump power may result in higher-order nonlinear effects; for example, Brillouin or Raman scattering Suppression of Raman and Brillouin scattering (for methods to mitigate such effects, see [28–30]). Therefore, we attenuated the signal to use 1W pump power. After the signal is coupled with a pump (ξ):

$$x_{n+1}^{(2)} = \sqrt{A}x_{n+1}^{(1)} + \xi \quad (3)$$

Then χ^3 -nonlinearity in dispersionless media will induce nonlinear phase shift, which is

proportional to the selected quadrature of the signal

$$x_{n+1}^{(3)} = x_{n+1}^{(2)} e^{i\gamma L |x_{n+1}^{(2)}|^2} \quad (4)$$

here, γ , L are nonlinear coefficients and fiber length, respectively. In general, the signal might be complex. By modulating the pump we can choose a quadrature of the desired signal. As an example, if one aims to process the real part of the signal, then the pump needs to be modulated as $\xi = 1$. Thus, in a limit of strong pump, in comparison to the signal, (which can be easily achieved by varying the attenuation factor), we obtain a sine transformation of the signal. The scheme is similar to [31, 32].

Overall, after nonlinear fiber, the output is

$$x_{n+1}^{(3)} = x_{n+1}^{(2)} e^{i\gamma L |x_{n+1}^{(2)}|^2}, \quad \Im x_{n+1}^{(3)} \approx \sin(\Im x_{n+1}^{(2)}), \quad (5)$$

For parameters: $\{\xi = 1, \gamma L |\xi|^2 = 2\pi, 2\sqrt{A}\xi\gamma L = 1\}$. Although the output is a complex signal, using a modulated strong pump makes it possible to receive output $F(x)$ dependent only on the real value of the input signal. Thus, although $x(n)$ is complex at each stage, it depends only on the real part of the signal at the previous stage and the real part of the input signal. Note that, here, all weight matrices W^{in} , W , W^{out} are real-valued. So, at the next stage, the imaginary part will be annihilated. While at the receiver where the states are collected, one selects only the real component for the matrix X . The resulting curve is plotted in Fig. 1(b), while the transformation from the numerical simulations of the standard off-shelf available highly nonlinear fiber [33], with nonlinear, dispersion, and attenuation coefficients of 10.5 1/W/km , -1.5 ps/nm/km , and 0.8 dB/km , respectively, results in the FESNA-transfer function shown in Fig. 1(b). Here and further in the paper, the numerical simulations were conducted using the standard split-step Fourier method (to ensure high precision we used a variable step size with maximum phase change 0.01). Note, to compensate for high losses in the fiber, the fiber length was optimized, which resulted in the relation $|\xi|^2\gamma L = 2.2\pi$. Overall, the effect of dispersion was negligible, due to low dispersion fiber and low signal power. Note that higher values of nonlinear coefficients are also available, including some twice as high [34].

2.3. Setup for dual-quadrature signal processing

In the case of dual-quadrature operation, we use the same concept of the modulated pump to select a quadrature and nonlinear fiber for nonlinear transformation (Fig. 1(d)). This can be achieved simultaneously by using the same piece of fiber in a loop configuration by coupling signals with the pumps before they undergo nonlinear transformation.

After mixing, the signal is

$$x_{n+1}^{(1)} = Wx_n + W^{in}u_{n+1} \quad (6)$$

in this case, matrices W , W^{in} and later W^{out} are complex. The signal is then attenuated $x_{n+1}^{(2)} = x_{n+1}^{(1)} \sqrt{A_1}$ and after a 3 dB coupler two copies of the signal are coupled with the pump. The pump is much stronger than the signal and modulated here as $\xi = 1 + i$:

$$x_{n+1}^{(3)CW} = x_{n+1}^{(2)}/\sqrt{2} + \xi; \quad x_{n+1}^{(3)CCW} = ix_{n+1}^{(2)}/\sqrt{2} + \xi \quad (7)$$

here CW and CCW stands for clockwise and counterclockwise. After the dispersionless nonlinear fiber, the output is

$$x_{n+1}^{(4)CW/CCW} = x_{n+1}^{(3)CW/CCW} e^{i\gamma L |x_{n+1}^{(3)CW/CCW}|^2}, \quad (8)$$

after propagation both signals are coupled back and attenuated

$$x_{n+1} = f(x) = \sqrt{A_2/2}(x_{n+1}^{(4)CW} + ix_{n+1}^{(4)CCW}) \quad (9)$$

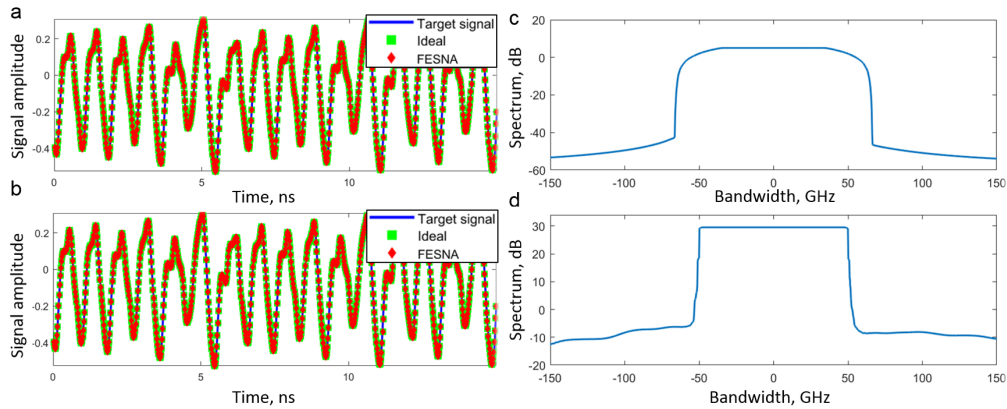


Fig. 2. FESNA for prediction tasks. The test signal is taken from a standard Mackey-Glass test, where we could perform prediction for a signal with 100 GHz bandwidth. For comparison, we plot (and as there is no significant difference visually, we show the mean square error (MSE)) the predicted signal obtained with \tanh -function (blue, solid, the mean squared error of test-regime $MSE = 4.8 \times 10^{-9}$), with \sin -function (green, squares, $MSE = 8 \times 10^{-7}$), FESNA-generated function (red, diamonds, $MSE = 6 \times 10^{-7}$). Panel (a) depicts one-step ahead prediction (i.e. for test-regime the previous target signal data points are used to generate further one-step ahead prediction) and panel (b) depicts feedback forcing or generation, (i.e. for test-regime the predicted data points are used to generate further predictions). One can see that, in the considered interval, both cases result in highly accurate performance comparable to that of ideal RC; bandwidth of (c) input signal and (d) output signal after FESNA.

here A, γ, L are attenuation and nonlinear coefficients and fiber length; also, for convenience, we chose $\gamma L = 2\pi$.

Although the output is a complex signal, using the aforementioned transformation one can receive sine-approximated transformation of both quadratures simultaneously, which is a close approximation of ideal sigmoid transformation. Thus, at each stage, both quadratures of the signal in the reservoir x_n are processed, while the collected states, signal X , are processed at the receiver side. We found that applying linear regression to augmented matrix $\tilde{X} = [XX^*]$ facilitates the training process and makes it possible to reduce the size of the reservoir. This way, the complex processing is conducted optically, while in the digital domain the complexity is reduced to a simple matrix multiplication.

3. Results

3.1. Application for prediction

To study the performance of the received RC, we used it for the prediction tasks - benchmarking task in machine learning is the Mackey-Glass chaotic time series. We use the Mackey-Glass delay equation

$$\frac{dy}{dt} = \frac{\alpha y(t - \tau)}{1 + y^\beta(t - \tau)} - \gamma y(t) \quad (10)$$

Following the example of [11] we chose parameters as $\tau = 17, \alpha = 0.2, \beta = 10, \gamma = 0.1$, an integration step of 0.17, downsampled to $t_s = 3$ to obtain the discrete time series used for processing. We used 1000 samples for training and 5000 for testing. The signal bandwidth was chosen to be 100 GHz. The performance comparison is plotted in Fig. 2. One can see that sigmoid, sine, and the generated functions give good performance. We compared the two

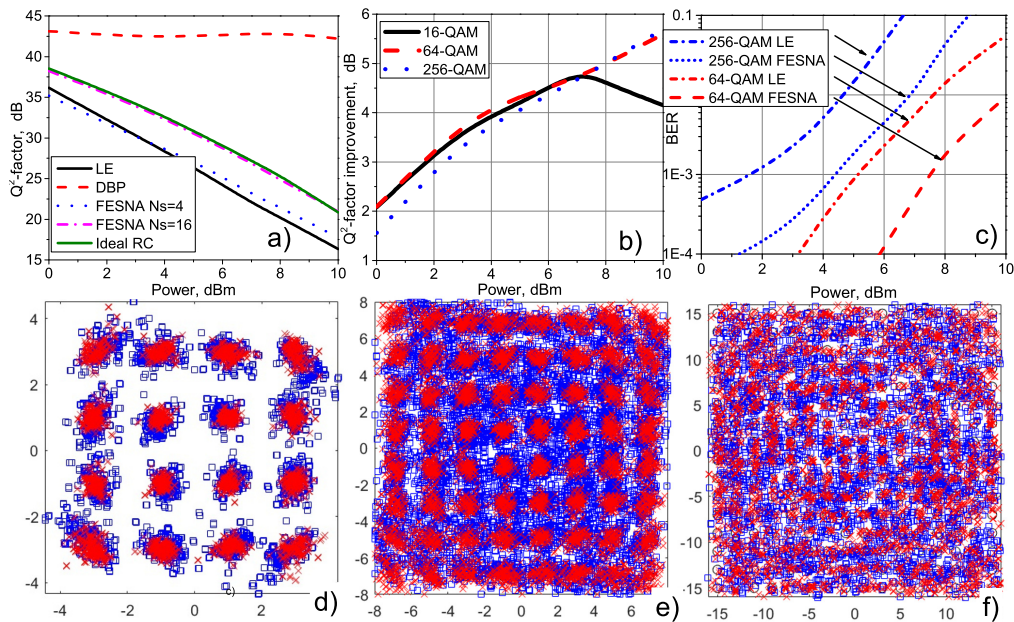


Fig. 3. Fiber reservoir computing for 16-, 64-, and 256-QAM signal processing. (a) Q^2 -factor for a signal with varied input power for a transmission distance of 100 km, processed with a linear equalizer (LE, which compensates dispersion and phase shift), digital back-propagation (DBP, with 16 samples per symbol and 50 steps per span), and fiber reservoir computing (FESNA, with four and 16 samples per symbol) with reservoir size 2^7 and training on 1000 symbols. Performance of ideal (sigmoid-based) RC is shown for comparison. (b) Q^2 -improvement due to FESNA-processing over linear equalization. The reservoir and signal parameters are the same as in Fig. 2 with a sampling rate of 16. (c) The corresponding BER. (d) 16-, (e) 64-, and (f) 256-QAM modulated signal after linear equalization (blue) and FESNA processing (red) for 10 dBm input power.

regimes and found that one-step prediction and feedback generation both demonstrate accurate performance in the considered interval.

3.2. Application for distortion mitigation in fiber-optic communication systems.

Next, we study the performance of the designed FESNA for compensating nonlinear signal distortions in fiber-optic communication systems. Performance of fiber-optic communication systems is limited by the efficiency and speed of employed signal processing [35], while new approaches are required beyond the state-of-the-art capabilities [36, 37]. We used a system with conventional parameters: a single 30 GBaud channel modulated with root-raised cosine pulses having 0.1 roll-off. Here we focus on single channel operation, leaving multichannel regime to further research. To examine the performance of FESNA for compensating deterministic nonlinear effects we study 100 km single span transmission with varied signal power (Fig. 3). The fiber parameters are 17 ps/nm/km dispersion and 1.4 1/W/km nonlinear coefficient. Our focus here is the nonlinear transmission regimes, so we use high input signal power.

We compare FESNA performance with DBP (50 steps per span and 16 samples per symbol) and linear equalization (LE), which compensates for the circular phase shift and other linear distortions. Note that digital back propagation is a standard technique for nonlinearity compensation in optical communication systems; particularly, in the absence of amplifier noise, it enables complete compensation of signal distortions, thus yielding the best performance. On the other

hand, linear equalization represents the minimum distortion compensation technique that is typically employed, thus presenting the lowest performance level compensation technique. We compensate dispersion before FESNA (this can be achieved by dispersion-compensating fiber or electronically). We use 1000 symbols for training and 2^{15} QAM symbols for testing, taking particular care to ensure random sequence of symbols [38]. To quantify the performance, we use Q^2 -factor ($Q^2 = 1/EVM^2$, where the error vector magnitude (EVM) is defined as in [39]).

First, we compare performance of FESNA to linear equalizer in Fig. 3(a) for 16-QAM signal. It can be seen that the performance of FESNA depends strongly on sampling rate. In particular, for a small sampling rate, such as 4, only linear effects are compensated. Increasing the sampling rate to 16 enables a significant improvement in Q^2 -factor, which grows with signal power.

We then examine compensation of nonlinear effects for higher order QAM. The performance comparison is plotted in Fig. 3(b), which shows that Q^2 -factor improvement is growing for higher nonlinearity. While the value of Q^2 -factor changes for various QAM constellations, the gain roughly remains the same until the nonlinearity is too high for FESNA and the gain decreases. The peak in Q^2 -factor improvement has higher tolerance for higher levels of modulation formats as they are more susceptible to nonlinearity, hence FESNA compensation results in higher gains. Fig. 3(c) shows the corresponding BERs, while the constellation diagrams without and with FESNA are depicted in Fig. 3(d)-(f) for 16-, 64-, and 256-QAM correspondingly. Overall, the figure illustrates that all-optical neural network realization based on FESNA enables high-efficiency mitigation of linear and nonlinear distortions.

4. Conclusions

In conclusion, we have proposed a new fiber-optic reservoir computing scheme that for the first time enables dual-quadrature signal processing with over 100 GHz bandwidth. The performance of the reservoir operation with high bandwidth was demonstrated on the standard Mackey-Glass prediction task. We also have shown that the designed FESNA scheme can be used for compensation of nonlinear distortions in fiber-optic transmission systems and can operate with various QAM signals. These presented results offer a novel type of photonic reservoir computing implementation with vast bandwidth range and dual-quadrature signal processing, which significantly broadens the application spectrum of technology.

Funding

Royal Academy of Engineering Research Fellowship RF\201718\17154.

References

1. M. A. Jarajreh, E. Giacoumidis, I. Aldaya, S. T. Le, A. Tsokanos, Z. Ghassemlooy, N. J. Doran, "Artificial neural network nonlinear equalizer for coherent optical OFDM," *IEEE Photon. Technol. Lett.* **27** 387-390 (2015).
2. E. Giacoumidis, S. T. Le, M. Ghanbarisabagh, M. McCarthy, I. Aldaya, S. Mhatli, M. A. Jarajreh, P. A. Haigh, N. J. Doran, A. D. Ellis, and B. J. Eggleton, "Fiber nonlinearity-induced penalty reduction in CO-OFDM by ANN-based nonlinear equalization," *Opt. Lett.* **P40**, 5113-5116 (2015).
3. M. Sorokina, S., Sygletos, and S. Turitsyn, "Sparse identification for nonlinear optical communication systems: SINO method," *Opt. Express* **24**, 30433-30443 (2016).
4. C. Hager and H. D. Pfister, "Nonlinear Interference Mitigation via Deep Neural Networks," *Optical Fiber Commun. Conf.*, paper W3A.4 (2018).
5. T. S. R. Shen and A. P. T. Lau, "Fiber nonlinearity compensation using extreme learning machine for DSP-based coherent communication systems," *OECC* 816-817 (2011).
6. S. Owaki and M. Nakamura, "Equalization of optical nonlinear waveform distortion using neural-network based digital signal processing," *Optoelectronics and Communications Conf. (Photon. Switching)* paper WA2-40 (2016).
7. M. Mahowald and R. Douglas, "A silicon neuron," *Nature* **354**, 515-518 (1991).
8. C. Eliasmith, T. C. Stewart, X. Choo, T. Bekolay, T. DeWolf, Y. Tang, D. Rasmussen, "A Large-Scale Model of the Functioning Brain," *Science* **338**, 1202-1205 (2012).
9. T. A. Engel, N. A. Steinmetz, M. A. Gieselmann, A. Thiele, T. Moore, K. Boahen, "Selective modulation of cortical state during spatial attention," *Science* **354**, 1140-1144 (2016).

10. A. N. Tait, T. F. Lima, E. Zhou, A. X. Wu, M. A. Nahmias, B. J. Shastri and P. R. Prucnal, "Neuromorphic photonic networks using silicon photonic weight banks," *Sci. Rep.* **7**, 7430(2017).
11. J. Bueno, D. Brunner, M. C. Soriano, and I. Fischer, "Conditions for reservoir computing performance using semiconductor lasers with delayed optical feedback," *Opt. Express* **25**, 2401-2412 (2017).
12. E. Cohen, D. Malka, A. Shemer, A. Shahmoon, Z. Zalevsky, and M. London, "Neural networks within multi-core optic fibers," *Sci. Rep.* **6**, 29080 (2016).
13. M. Lukosevicius and H. Jaeger, "Reservoir Computing Approaches to Recurrent Neural Network Training," *Computer Science Review* **3**, 127-149 (2009).
14. M. Bauduin, S. Massar, and F. Horlin, "Non-linear satellite channel equalization based on a low complexity echo state network," *Information Science and Systems Conference*, 99-104 (2016).
15. K. Vandoorne, J. Dambre, D. Verstraeten, B. Schrauwent, and P. Bienstman, "Parallel reservoir computing using optical amplifiers," *IEEE Trans. Neural Netw.* **22**, 1469-1481 (2011).
16. A. Argyris, J. Bueno, I. Fischer, "Photonic machine learning implementation for signal recovery in optical communications," *Sci. Rep.* **8** (1), 8487 (2018).
17. J. Bueno, S. Maktoobi, L. Froehly, I. Fischer, M. Jacquot, L. Larger, and D. Brunner, "Reinforcement Learning in a large scale photonic Recurrent Neural Network," *Optica (OSA)* **5**, 756-760 (2018).
18. <https://www.technologyreview.com/s/404874/smart-fibers/>
19. J. N. Kutz, X. Fu, and S. Brunton, "Self-tuning fiber lasers: machine learning applied to optical systems," *Advanced Photonics Conf. paper NTu4A.7* . (2014).
20. T. Baumeister, S. L. Brunton, and N. J. Kutz, "Deep learning and model predictive control for self-tuning mode-locked lasers," *J. Opt. Soc. Am. B* **35**, 617-626 (2018)
21. Q. Ling, Z. Gu, and K. Gao, "Smart design of a long-period fiber grating refractive index sensor based on dual-peak resonance near the phase-matching turning point," *Appl. Opt.* **57**, 2693-2697 (2018).
22. <http://www.optonicus.com/infoco/>
23. L. Appeltant, M.C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C.R. Mirasso and I. Fischer, "Information processing using a single dynamical node as complex system," *Nat. Commun.* **2**, 468 (2011).
24. F. Dupont, A. Smerieri, A. Akrouf, M. Haelterman, and S. Massar, "Fully analogue photonic reservoir computer," *Sci. Rep.* **6**, 22381 (2016).
25. Y. Paquot, F. Dupont, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman and S. Massar, "Optoelectronic reservoir computing," *Sci. Rep.* **2**, 287 (2012).
26. L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutierrez, L. Pesquera, C. R. Mirasso, and I. Fischer, "Photonic information processing beyond Turing: an optoelectronic implementation of reservoir computing," *Opt. Express* **20**, 3241-3249 (2012).
27. H. Jaeger, "Adaptive nonlinear system identification with echo state networks," *Advances in Neural Information Processing Systems* **15**, 593-600 (2003).
28. A. Hideur, T. Chartier, C. Oezkul, and F. Sanchez, "Dynamics and stabilization of a high power side-pumped Yb-doped double-clad fiber laser," *Opt. Commun.* **186**, 311-317 (2000).
29. M. Heck, V. Bock, R. G. Kraemer, D. Richter, T. A. Goebel, C. Matzdorf, A. Liem, T. Schreiber, A. TÄijnnemann, S. Nolte, "Mitigation of stimulated Raman scattering in high power fiber lasers using transmission gratings," *Fiber Lasers XV: Technology and Systems* **10512**, 105-121 (2018).
30. C. L. Guinrand, "Stimulated Brillouin scattering threshold variations due to bend-induced birefringence in a non-polarization-maintaining fiber amplifier," *CLEO JW2A-23* (2014).
31. M. S. Sorokina and S. K. Turitsyn, "Regeneration limit of classical Shannon capacity," *Nat. Commun.* **5**, 3861 (2014).
32. M. Sorokina, S. Sygletos, and S. Turitsyn, "Regenerative Fourier transformation for dual-quadrature regeneration of multilevel rectangular QAM," *Opt. Lett.* **40**, 3117-3120 (2015).
33. <https://fiber-optic-catalog.ofsoptics.com/item/optical-fibers/highly-nonlinear-fiber-optical-fibers1/hnlf-pm-highly-non-linear-fiber-modules>
34. T. Okuno, M. Onishi, T. Kashiwada, S. Ishikawa, and M. Nishimura, "Silica-based functional fibers with enhanced nonlinearity and their applications," *IEEE J. Sel. Top. Quantum Electron.* **5** 1385-1391 (1999).
35. A. D. Ellis, M. E. McCarthy, M. A. Z. Al Khateeb, M. Sorokina, and N. J. Doran, "Performance limits in optical communications due to fiber nonlinearity," *Adv. Opt. Photon.* **9**, 429-503 (2017).
36. M. Sorokina, S. Sygletos, and S. Turitsyn, "Ripple distribution for nonlinear fiber-optic channels," *Opt. Express* **25**, 2228-2238 (2017)
37. E. Agrell and M. Secondini, "Information-theoretic tools for optical communications engineers," 2018 IEEE Photonics Conference (IPC) (2018).
38. T. A. Eriksson, H. Buelow, A. Leven, "Applying neural networks in optical communication systems: possible pitfalls," *IEEE Phot. Technol. Lett.* **29**, 2091-2094 (2017).
39. W. Freude, R. Schmogrow, B. Nebendahl, M. Winter, A. Josten, D. Hillerkuss, S. Koenig, J. Meyer, M. Dreschmann, M. Huebner, C. Koos, J. Becker, J. Leuthold, "Quality metrics for optical signals: eye diagram, Q-factor, OSNR, EVM and BER," *International Conference on Transparent Optical Networks* (2012).