Multidimensional fiber echo state network analogue

IOP Publishing J. Phys. Photonics 2 (2020) 044006 https://doi.org/10.1088/2515-7647/abb584

Journal of Physics: Photonics

PAPER

Multidimensional fiber echo state network analogue

Mariia Sorokina

Aston University, Aston Institute of Photonic Technologies, B4 7ET, United Kingdom

E-mail: m.sorokina@aston.ac.uk

Keywords: neuromorphic computing, fiber-optics, optical signal processing

OPEN ACCESS

RECEIVED
3 June 2020

REVISED
29 July 2020

ACCEPTED FOR PUBLICATION
4 September 2020

PUBLISHED
1 October 2020

Abstract

Optical neuoromorphic technologies enable neural network-based signal processing through a specifically designed hardware and may confer advantages in speed and energy. However, the advances of such technologies in bandwidth and/or dimensionality are often limited by the constraints of the underlying material. Optical fiber presents a well-studied low-cost solution with unique advantages for low-loss high-speed signal processing. The fiber echo state network analogue (FESNA), fiber-based neuromorphic processor, has been the first technology suitable for multichannel high bandwidth (including THz) and dual-quadrature signal processing. Here we propose the multidimensional FESNA (MD-FESNA) processing by utilizing multi-mode fiber non-linearity. Thus, the developed MD-FESNA is the first neuromorphic technology which augments all aforementioned advantages of FESNA with multidimensional spatio-temporal processing. We demonstrate the performance and flexibility of the technology on the example of prediction tasks for hyperchaotic systems. These results will pave the way for a high-speed neuromorphic processing of multidimensional tasks, hardware for spatio-temporal neural networks and open new application venues for fiber-based spatio-temporal multiplexing.

1. Introduction

Optical neuromorphic computing has attracted increasing attention over the past decades [1, 2]. An important advantage of the technology is its ability to execute a neural network (NN) in an optical domain providing a significant enhancement in speed. Among variety of approaches [3–5], reservoir computing (RC) architecture enables an optical NN implementation with relaxed training requirements. In particular, echo state network (ESN) [6], a type of RC, has been realised via various computing platforms [7, 8]. At the same time there is a surge in developments aimed at increasing operational bandwidth of the neuromorphic systems. Recently an optical ESN utilizing semiconductor laser [9] and silicon photonics for reservoir technology [10] have been developed for processing 20 GHz and 40 GHz bandwidth signals correspondingly. While recently we introduced a fiber-based neuromorphic technology - fiber ESN analogue (FESNA) [11], which realizes signal mixing and neural response by utilizing inherent fiber properties: dispersion and the Kerr non-linearity - this enabled to achieve a dual-quadrature [11], high bandwidth (including THz) [12, 13] and multi-channel [14] neuromorphic signal processing for the first time.

On the other hand, many applications in data analysis and machine learning, such as object recognition or decision making, require processing of multidimensional data. The increased complexity of analysis also bears immediate effect on speed, energy efficiency and complexity of the processor. It is beneficial therefore to incorporate more physical dimensions into the related neuromorphic hardware. Indeed, these are of high importance for many applications including optical communications, emerging 6G and Internet of Things (IoT), where space division multiplexing (SDM) is employed for enhancing the transmission rate requiring multi-dimensional signal equalization for compensation of non-linear impairments [15–18] and smart fiber-based technologies, including lasers and sensors [19–22].

Moreover, many complex dynamical effects and systems can be analysed only by processing all dimensions simultaneously. For example, hyperchaos [23, 24], which occurs only in multidimensional

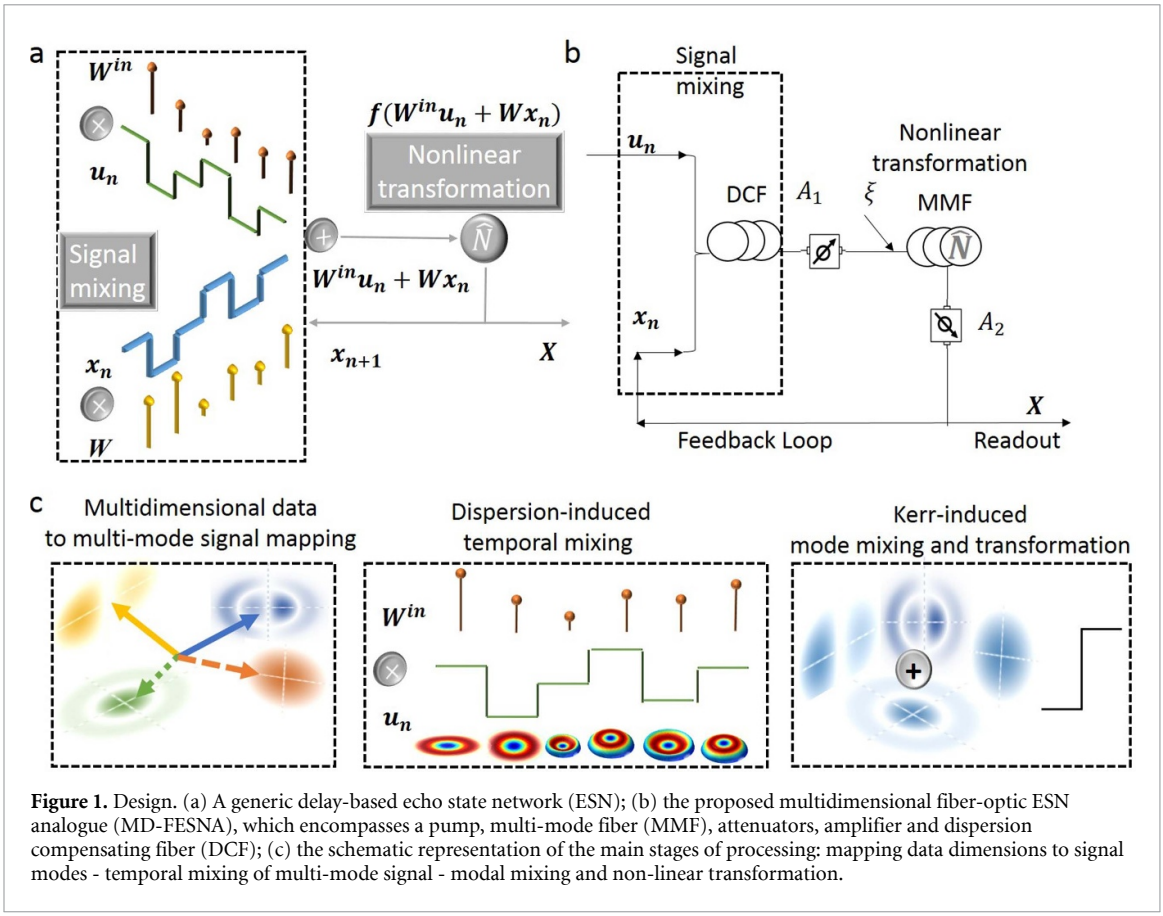

Figure 1. Design. (a) A generic delay-based echo state network (ESN); (b) the proposed multidimensional fiber-optic ESN analogue (MD-FESNA), which encompasses a pump, multi-mode fiber (MMF), attenuators, amplifier and dispersion compensating fiber (DCF); (c) the schematic representation of the main stages of processing: mapping data dimensions to signal modes - temporal mixing of multi-mode signal - modal mixing and non-linear transformation.

systems. Furthermore, due to the increasing interest in the studies of spatio-temporal NNs there is an interest in spatio-temporal neuromorphic platforms.

The developments of fibers with multiple spatial modes and with engineered properties on demand [15, 16, 25] enable to apply multi-mode fibers for neuromorphic computing. In particular, multi-core fiber-based neuromorphic technology has already been demonstrated, where cores were utilized as virtual neurons and the strength of synaptic interactions was varied by amplifiers [26].

Here we are focused on a different task – incorporating spatial fiber modes (similarly one can use cores) to process all signal dimensions simultaneously, which enables to solve complex machine learning problems. Similarly to FESNA [11], we use temporal multiplexing for realizing virtual neurons and augment it further by spatial multiplexing to realize multidimensional signal interactions.

We show that by incorporating multi-mode fiber (MMF), we can achieve multidimensional neuromorphic processing - multidimensional FESNA (MD-FESNA). This will be beneficial for solving complex multidimensional tasks, which is demonstrated here numerically on the example of hyperchaotic systems. Moreover, MD-FESNA incorporates spatio-temporal dimensions - crucial for many effects and applications, including spatio-temporal NNs. Thus, MD-FESNA is the first multi-dimensional spatio-temporal neuromorphic technology.

2. Design

Here we describe a design of optical neuromorphic architecture based on a multi-mode fiber for multidimensional signal processing. The design is based on the ESN as it enables NN realisation with one non-linear element and a feedback loop (see figure 1(a)). The typical ESN comprises the following stages of signal processing:

(i) signal mixing through random weight matrices $\mathbf{W}^{in}$ and $\mathbf{W}$, which are multiplied to the input $\mathbf{u}_n$ and feedback $\mathbf{x}_n$ signals correspondingly: $\mathbf{x}_{n+1}^{(i)} = \mathbf{W}^{in}\mathbf{u}_n + \mathbf{W}\mathbf{x}_n$. In multidimensional case matrices incorporate mixing among all signal dimensions.

(ii) non-linear transformation: $\mathbf{x}_n^{(ii)} = f(\mathbf{W}^{in}\mathbf{u}_n + \mathbf{W}\mathbf{x}_n)$, which is usually a sigmoid or tanh function. The resulted signal is then fed back for the subsequent mixing and the copy is also collected at the receiver $\mathbf{X}$.

The state matrix is then used to obtain the optimal output weight matrix $\mathbf{W}^{out}$ through linear regression: $\mathbf{y} = \mathbf{W}^{out}\mathbf{X}$. Thus, the only optimized element is the output weight matrix, optimization of which requires a

linear regression only. This property is an important advantage of the ESN compared to other NNs, training of which requires complex optimization algorithms.

Since the first design of the fiber ESN analogue (FESNA) [11], where only non-linear transformation was realised optically, we have shown that both signal mixing and non-linear transformation can be realised optically using fiber dispersion and non-linearity correspondingly (DM-FESNA) [13]. Unlike conventional neuromorphic systems, here we assume an optical input signal u to be multidimensional: a result of multidimensional digital data encoded to the analog multi-mode signal [15, 16, 27, 28] or multi-mode signal originated from communication systems (e.g. SDM). As shown in figure 1(c), each input data dimension or signal mode is mapped to the corresponding mode of the MMF (in the illustration there are 4 data/signal dimensions and 4 MMF modes, correspondingly).

One of the advantages of utilizing fiber properties for both stages (i)–(ii) is that it enables high bandwidth signal processing [13]. Moreover such design enables the first multichannel neuromorphic processing, which is of high importance for many applications, in particular optical communications. Here we demonstrate that augmented by multi-mode fiber we can achieve simultaneous multidimensional signal processing - multidimensional FESNA (MD-FESNA) (see figure 1(b)). Moreover we can significantly reduce the complexity of previous realisations, as one can realise MD-FESNA using one spool of fiber and only one pump. The setup follows the stages outlined above for ESN:

(i) To achieve signal mixing we utilize fiber dispersion:

$$\hat{\mathbf{D}}[u(t)] = \sqrt{\frac{i}{2\pi D}} \int dt' u(t') \exp\Big(- i\frac{(t-t')^2}{2D}\Big). \tag{1}$$

To control the amount and strength of mixing one can change the accumulated dispersion D of the dispersion compensating fiber (DCF). Dispersion induces temporal mixing between signal samples, each of which is a multidimensional signal being a combination of the modes (see figure 1(c)).

(ii) To realize non-linear transformation we use the inherent Kerr-non-linearity of the fiber, which in multi-mode case is described by the Manakov equation for the M-modes [29–31], so that each mode $p = 1..M$ is transformed as follows:

$$\frac{\partial U_p}{\partial z} = -\frac{\alpha}{2}U_p - \beta_1^p \frac{\partial U_p}{\partial t} - i\frac{\beta_2^p}{2}\frac{\partial^2 U_p}{\partial^2 t} + i\kappa_{pp}\gamma_{pp}|U_p|^2 U_p + i\kappa_{pq}\sum_{q\neq p}^{M}\gamma_{pq}|U_q|^2 U_p, \tag{2}$$

here deterministic distortions are described by fiber losses α, second-order dispersion β_2 and non-linearity coefficient γ. The formula captures both a) weak and b) strong coupling regimes, when a) $\kappa_{pp} = 8/9, \kappa_{pq} = 4/3$ and b) $\kappa = \kappa_{pq} = \kappa_{pq} = 8\,\mathrm{M}/3/(2\,\mathrm{M}+1)$. As an example, further we use strong coupling.

The resulted effect can be explained as follows. The Kerr-non-linearity induces a non-linear phase shift:

$$x^{(ii)} = F_{MMF}[x^{(i)})] = x^{(i)} e^{i\gamma L\kappa ||x^{(i)}||^2} \tag{3}$$

(above we assumed no dispersion for explanation purposes only). By coupling the signal with the pump one can achieve signal-pump beating, which in the limit of strong pump (compared to signal) $\mathbf{x}^{(i)} + \xi$ results in harmonic oscillations of $\mathbf{x}^{(ii)}$. In the previous FESNA designs we used parameters and loop mirror design to achieve sine-transformation with the aim to approximate tanh function, typically used for the ESN. However, in most cases this is not necessary and, as we will show further, there is a broad flexibility in the choice of setup parameters. Thus, the Kerr non-linearity results in the mode mixing and non-linear transformation, the properties of which can be governed by the pump and fiber parameters (see figure 1(b, c)).

In MD-FESNA, unlike the ESN design above, we do not require mixing of modes in the signal mixing stage (thus, reducing dimensionality of the weight matrices to mixing of the signal samples in time only) as the signal modes will mix at the non-linear stage naturally through non-linearity.

3. Performance and discussion

To illustrate the performance of the proposed setup we test it on a prediction task for the multidimensional chaotic systems. In particular, we choose to consider hyperchaotic systems as they demonstrate the impact of mutually interconnected dimensions, which require to be processed simultaneously. In particular, we use a well-studied hyperchaotic system - 4D Rössler system [23, 24]:

$$\dot{x} = -y - z; \quad \dot{y} = x + ay + w; \quad \dot{z} = b + xz; \quad \dot{w} = -cz + dw. \tag{4}$$

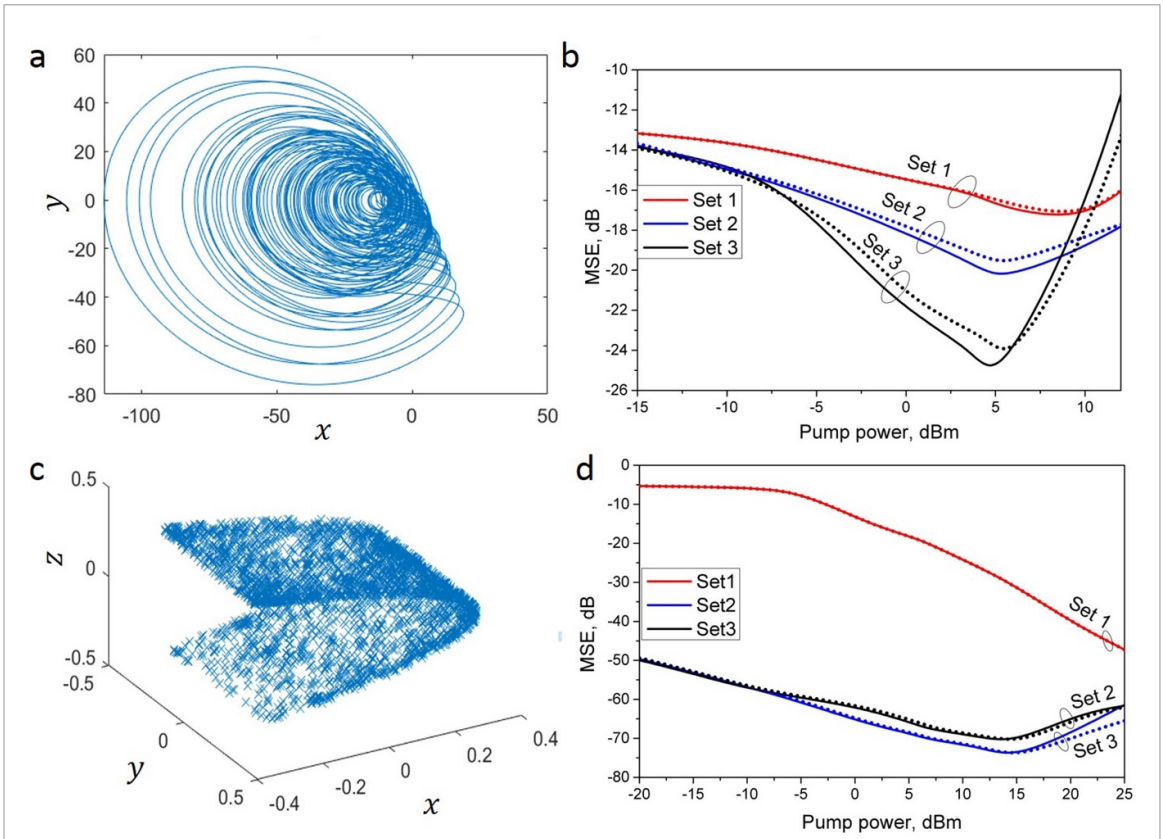

Figure 2. Performance results. (a) Plane projection of the hyperchaotic 4D Rössler system; (b) The mean squared error (MSE) for MD-FESNA processor with varied set of setup parameters from table 1, with parameters for HNLF and standard MMF shown by solid and dotted curves correspondingly; (c) The hyperchaotic 3D Hénon map; (d) MSE for MD-FESNA processor with parameters given in table 1.

Table 1. MD-FESNA parameters, including standard optical communications and HNLF (the latter is shown in brackets) MMF fibers.

Parameter	Value
Maximum phase change	0.01
MMF Attenuation coefficient	0.2 (0.8) dB/km
MMF CD	17 (1.5) ps/nm/km
MMF Nonlinearity	1.4 (10.5) 1/W/km
DCF Dispersion	$D = 20ps^2$
Gain	$G = 6$ dB
Feedback attenuator	$A_2 = -20$ dB

	Set1	Set2	Set3
MMF Length	50 (5) m	500 (50) m	750 (75) m
Reservoir attenuator	$A_1 = -24.3$	$A_1 = -18.7$	$A_1 = -20$ dB

The plane projection is plotted in figure 2(a) for a standard set of parameters: $a = 0.25, b = 3, c = 0.5, d = 0.05$ and initial conditions: $x_0 = -10; y_0 = 6; z_0 = 0; w_0 = 10$. The equations are simulated by the Runge-Kutta method with step 10^{-4}, then downsampled with sampling rate 16 and normalized to average power 0.05 to form the data for processing. Here we utilize two signal quadratures (real and imaginary) and two modes. For benchmarking we compare the performance to that of a standard ESN architecture with four dimensions and following parameters: reservoir size $N = 256$, number of training and testing samples $N_{tr} = 1000$ and $N_{test} = 2000$ with tanh activation function and weights chosen randomly from the intervals $[-0.5...0.5]$ and $[-0.5i...0.5i]$ for all quadratures and modes. We evaluate the performance in terms of the mean squared error (MSE) normalized to the average signal power. The maximum MSE after ESN processing is $MSE_{ESN} = -20.5$ dB compared to $MSE_{LR} = -8.5$ dB in case of a standard ridge linear regression. Note, that in the ESN the mode mixing in matrix **W** is essential, without which the performance is comparable to that of linear regression, which highlights that processing data dimensions as separate data streams is equivalent to linear regression and multidimensional approach is essential.

The parameters of MD-FESNA are chosen according to the table 1. Here we consider two examples of multi-mode fiber (MMF): standard MMF fiber used in optical communication networks and systems [32, 33] and highly non-linear fiber (HNLF) [34, 35], parameters for the latter are given in brackets alongside the corresponding parameters for a standard MMF. The HNLF offers higher values of non-linearity (resulting in shorter fiber length), yet they have higher values of attenuation. The corresponding MSE is plotted in figure 2(b) as a function of varied pump power and for different sets of MMF parameters. The HNLF and standard optical communications fiber scenarios demonstrate similar performance. While in the latter case higher pump powers are required to achieve peak performance, which is due to longer fiber lengths (see Sets2,3) and the achieved minimum MSE is bigger than in HNLF case, which is due to additional dispersion-induced signal-signal mixing, which manifests itself for longer fiber lengths (see Sets2,3). One can see that even in case of smaller non-linearity (Set1) one can achieve high performance compared to linear regression, while Set 2 gives result comparable to that of the ESN. The increase of the MMF length (Set3) results in further improvement of the performance surpassing that of the ESN with tanh activation function. This could be explained in terms of a choice of activation function, as while tanh function is commonly used in the ESN and other NNs, nevertheless, it is not necessarily the optimum function. The optimization of pump power shows that the optimum performance is achieved around the pump power 5 dBm with the corresponding MSE significantly surpassing MSE resulted from linear regression.

Next we compare the performance of the proposed setup with the same set of parameters for the prediction task based on the 3D hyperchaotic Hénon map [36]:

$$x_{n+1} = a - y_n^2 - bz_n; \quad y_{n+1} = x_n; \quad z_{n+1} = y_n. \tag{5}$$

The plane projection is plotted in figure 2(c) for a standard set of parameters: $a = 1.9, b = 0.03$ and initial conditions were chosen as: $x_0 = 0.1; y_0 = 0.2; z_0 = -0.1$. The chaotic 2D Hénon map is often used for benchmarking in neuromorphic systems [37]. Here we utilize three modes and a single signal quadrature – signal amplitude. We compare the performance to that of the standard ESN architecture with three dimensions and the same parameters as in the above example. The maximum MSE after the ESN processing is $MSE_{ESN} = -59$ dB, while linear regression returns $MSE_{LR} = -0.9$ dB.

The parameters of MD-FESNA are also kept the same as in the previous example, see table 1, which allows us to study the flexibility of the setup for solving different tasks when parameters are fixed. The pump power is expected to increase as it represents the total power distributed over the increased number of modes. The corresponding MSE, plotted in figure 2(d) as a function of varied pump power, demonstrates that by using the same set of parameters (for example Set2 or Set3), one can achieve good performance for various tasks. Here the difference in performance for highly non-linear and standard MMF is less pronounced than in the previous example (see figure 2(b)), processing of which required less non-linearity yet was more sensitive to optimization.

4. Conclusions

In summary, we have proposed a fiber-based technology for dual-quadrature high bandwidth multidimensional neuromorphic signal processing. The setup enables to incorporate spatial and temporal dimensions enabling high speed performance. We demonstrated an application of the technology for prediction of multidimensional hyperchaotic systems. The design is flexible and applicable to various tasks with the fixed set of parameters. The technology, based on the low-cost off-the-shelf fibers, offers unique advantages such as high bandwidth and multidimensionality and opens new opportunities for advanced fiber-based signal processing addressing variety of applications ranging from optical communications and 5/6G to lasers and sensors.

Acknowledgment

This project was supported by the Royal Academy of Engineering under the Research Fellowship scheme *RF*/201718/17154.

ORCID iD

Mariia Sorokina https://orcid.org/0000-0001-6082-0316

References

[1] Prucnal P R and Shastri B J 2017 *Neuromorphic Photonics* (Boca Raton, FL: CRC Press)

[2] Lugnan A, Katumba1 A, Laporte F, Freiberger M, Sackesyn S, Ma C, Gooskens E, Dambre J and Bienstman P 2020 Photonic neuromorphic information processing and reservoir computing *APL Photonics* **5** 020901
[3] Eliasmith C, Stewart T C, Choo X, Bekolay T, DeWolf T, Tang Y and Rasmussen D 2012 A Large-Scale Model of the Functioning Brain *Science* **338** 1202–5
[4] Engel T A, Steinmetz N A, Gieselmann M A, Thiele A, Moore T and Boahen K 2016 Selective modulation of cortical state during spatial attention *Science* **354** 1140–4
[5] Tait A N, Lima T F, Zhou E, Wu A X, Nahmias M A, Shastri B J and Prucnal P R 2017 Neuromorphic photonic networks using silicon photonic weight banks *Sci. Rep.* **7** 7430
[6] Jaeger H 2001 The 'echo state' approach to analysing and training recurrent neural networks GMD Report 148 GMD - German National Research Institute for Computer Science
[7] 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
[8] Vandoorne K, Mechet P, Van Vaerenbergh T, Fiers M, Morthier G, Verstraeten D, Schrauwen B, Dambre J and Bienstman P 2014 Experimental demonstration of reservoir computing on a silicon photonics chip *Nat. Commun.* **5** 3541
[9] Argyris A, Bueno J and Fischer I 2018 Photonic machine learning implementation for signal recovery in optical communications *Sci. Rep.* **8** 8487
[10] Katumba A, Yin X, Dambre J and Bienstman P 2019 A neuromorphic silicon photonics nonlinear equalizer for optical communications with intensity modulation and direct detection *J. Lightwave Technol.* **37** 2232–9
[11] Sorokina M, Sergeyev S and Turitsyn S 2019 Fiber Echo State Network Analogue for High-Bandwidth Dual-Quadrature Signal Processing *Opt. Express* **27** 2387–95
[12] Sorokina M 2019 High bandwidth all-optical fiber-based neuromorphic signal processing *Conf. on Optical Communications* p P39
[13] Sorokina M 2020 Dispersion-managed fiber echo state network analogue with high (including THz) bandwidth *J. Lightwave Technol.* **38** 3209–13
[14] Sorokina M 2020 Multi-channel optical neuromorphic processor for frequency-multiplexed signals submitted DOI: https://doi.org/10.13140/RG.2.2.33164.46720
[15] Fontaine N K *et al* 2015 30 × 30 MIMO transmission over 15 spatial modes *in Optical Fiber Conf. Post Deadline Papers* p paper Th5C.1. OSA Technical Digest (online) (Optical Society of America)
[16] Winzer P J 2015 Scaling optical fiber networks: challenges and solutions *Opt. Photonics News* **26** 28–35
[17] Vázquez C, Montero D S, Ponce W, Lallana P C and Larrabeiti D 2016 Multimode fibers in millimeter-wave evolution for 5G cellular networks *Proc. SPIE 9772 Broadband Access Communication Technologies X* p 97720F February 12
[18] Kasmia M, Mhatlia S, Bahloulab F, Dayoubc I and Ohd K 2020 Performance analysis of UFMC waveform in graded index fiber for 5G communications and beyond *Opt. Commun.* **454** 124360
[19] Fang S, Li B, Song D, Zhang J, Sun W and Yuan L 2013 A smart graded-index multimode fiber based sensor unit for multi-parameter sensing applications *Opt. Photon. J.* **03** 265–7
[20] (https://www.technologyreview.com/2005/11/08/101125/smart-fibers/)
[21] Picozzi A, Millot G and Wabnitz S 2015 Nonlinear virtues of multimode fibre *Nat. Photon.* **9** 289–91
[22] Wright L G, Christodoulides D N and Wise F W 2015 Controllable spatiotemporal nonlinear effects in multimode fibres *Nature Photon.* **9** 306–10
[23] Rössler O E 1979 An equation for hyperchaos *Phys. Lett.* **71** 155–7
[24] Letellier C and Rössler O E 2007 Hyperchaos *Scholarpedia* **2** 81936
[25] Amphawan A and Masunda T 2020 Mode generation in elliptical core few mode fiber *Conf. Proc.* vol 2203 p 020008
[26] Cohen E, Malka D, Shemer A, Shahmoon A, Zalevsky Z and London M 2016 Neural networks within multi-core optic fibers *Sci. Rep.* **6** 29080
[27] Jung Y, Kang Q, Alam S and Richardson D J 2020 Spatial Multiplexing: Technology chapter 8 *Optical Communication Systems Limits and Possibilities* ed Ellis A and Sorokina M (Singapore: Jenny Stanford Publishing)
[28] Birks T A, Gris-Sánchez I, Yerolatsitis S, Leon-Saval S G and Thomson R R 2015 The photonic lantern *Adv. Opt. Photon.* **7** 107–67
[29] Mumtaz S, Essiambre R-J and Agrawal G P 2013 Nonlinear propagation in multimode and multicore fibers: generalization of the Manakov equations *J. Lightwave Technol.* **31** 398–406
[30] Mecozzi A, Antonelli C and Shtaif M 2012 Nonlinear propagation in multi-mode fibers in the strong coupling regime *Opt. Express* **20** 11673–8
[31] Xiao Y, Essiambre R J, Desgroseilliers M, Tulino A M, Ryf R, Mumtaz S and Agrawal G P 2014 Theory of intermodal four-wave mixing with random linear mode coupling in few-mode fibers *Opt. Express* **22** 32039–59
[32] (https://www.furukawa.co.jp/en/product/catalogue/)
[33] (https://fiber-optic-catalog.ofsoptics.com/item/optical-fibers/highly-nonlinear-fiber-optical-fibers1/hnlf-pm-highly-non-linear-fiber-modules)
[34] Mumtaz S, Essiambre R-J and Agrawal G P 2013 Nonlinear Propagation in Multimode and Multicore Fibers: Generalization of the Manakov Equations *J. Lightwave Technol.* **31** 398–406
[35] Sorokina M, Sygletos S and Turitsyn S 2016 Sparse identification for nonlinear optical communication systems: SINO method *Opt. Express* **24** 30433–43
[36] Baier G and Klein M 1990 Maximum hyperchaos in generalized Hénon map *Phys. Lett.* **151** 281–4
[37] Rodan A and Tino P 2011 Minimum complexity echo state network *IEEE T. Neural Netw.* **22** 131–44