Dispersion-managed fiber echo state network analogue with high (including THz) bandwidth

Mariia Sorokina

***Abstract*—We propose a design for high (including THz) bandwidth neuromorphic signal processing based on fiber echo state network analogue and demonstrate through numerical modeling its efficiency for distortion mitigation in optical communications with high-order (including 1024-QAM) formats. This is achieved via all-optical implementation of masking and neural functionality by utilizing dispersion and nonlinear properties of the fiber. The design is flexible and format-transparent making it relevant for future communication systems. The model is simple to implement, as it requires only two attenuators and pumps in addition to a traditional nonlinear optical loop mirror (NOLM), and offers a vast range of optimization parameters.**

***Index Terms*—Neuromorphic computing, signal processing, Optical communications**

I. Introduction

The demand for faster realisation of complex signal processing algorithms drives development of novel optical signal processing approaches. Optical neuromorphic processing, specifically Echo State Networks (ESN) [1], enables optical realisation of neural networks with relaxed training complexity. One of the first demonstrations of an optical ESN utilizing GHz bandwidth was based on a semiconductor laser, which was also used recently for pulse amplitude modulated (PAM) signal processing [2]. Recently we introduced a new approach - fiber ESN analogue (FESNA) [3], which makes it possible to receive nonlinear "neuron" functionality via signal-pump beating. FESNA has enabled high bandwidth (over 30 GHz) and dual-quadrature signal processing. However, the bandwidth and processing speed were limited by the electrical component, which realised synaptic weights. To overcome this limit, we proposed to utilize fiber dispersion for synaptic signal mixing [4] - dispersion-managed FESNA (DM-FESNA). This was an important step towards implementation of neuromorphic technology for signal processing: while many signal processing applications require feasibility in bandwidth range and multidimensional processing (such as multichannel signals in optical communications), conventional neuromorphic technologies are limited to a few GHz bandwidth and a single channel operation due to fundamental limitations of the utilized materials, such as silicon or semiconductor. On the other hand, modern transmission systems operate at the bandwidth range of a few THz with ongoing research on ultra high bandwidth communications (incorporating 10 THz and higher).

Here we elaborate on the design of DM-FESNA and illustrate its flexibility with demonstrations ranging from 30 GHz to 1 THz. We demonstrate the application of the proposed system on the example of distortion mitigation tasks of high-order quadrature amplitude modulated (QAM) signals. The design is flexible which is demonstrated by utilizing the same setup with fixed parameters for processing signals ranging from 30 GHz to 1 THz bandwidth modulated with formats from 64- to 1024-QAM. Although the demonstration is limited to a single channel operation, together with the ongoing research of multichannel DM-FESNA [5], the work paves the way to further exploration and exploitation of fiber-based neuromorphic technology resolving limitations of other platforms.

Aston University, B4 7ET, UK

II. Method

Typical neural network consists of virtual nodes, which emulate neurons, by mixing signals from neighboring nodes with the given weights and performing nonlinear processing, see Fig. 1(a). While neural networks usually require a significant amount of training, Echo state network (ESN) relaxes this complexity as it assumes nodes in network to be randomly connected and training involves only optimization of the output weights. Moreover, ESN can be realised optically by a single nonlinear element with a feedback delay loop [6].

One can distinguish two major parts in ESN processing: signal mixing and nonlinear transformation, denoted in Fig. 1(b) by dashed square and grey circle correspondingly. Their optical analogues are depicted in Fig. 1(c), while the impact of each component is represented in Fig. 1(d) demonstrating its relation to the ESN concept in Fig. 1(b). At the first stage, the input signal $\mathbf{u}$ is injected gradually to the setup, as denoted by the sequence $\mathbf{u}_n$, and the samples are constantly mixed together as defined by an input weights matrix $\mathbf{W}^{in}$ and with a signal from a feedback loop $\mathbf{x}_n$ through a different weight matrix $\mathbf{W}$. This signal mixing process is emulating synaptic interactions, see dashed square in Fig. 1(b).

At the second stage, the nonlinear response of the neuron is modeled by the nonlinear activation function, denoted by a circle in Fig. 1(b). The output is then fed back to the loop.

Meanwhile, the output signals from the feedback loop are gradually collected at the receiver $\mathbf{X} = [...\mathbf{x}_n, \mathbf{x}_{n+1}, ...]$. Creating an augmented state matrix. For each sample of the input signal $\mathbf{u}_n$, N-dimensional $\mathbf{x}_n$-vector is collected. At the final stage the state matrix is used to obtain the optimum weights $\mathbf{W}^{out}$ by a simple linear regression or pseudo-inversion; this can be done offline or online. Thus, one can achieve neural network processing in optical domain.

Fiber-optic realisation of ESN - fiber ESN analogue (FESNA) - was proposed in [3], where fiber Kerr nonlinearity was used to emulate neural response, denoted by a circle in Fig. 1(b-c). The scheme enabled to process signals with

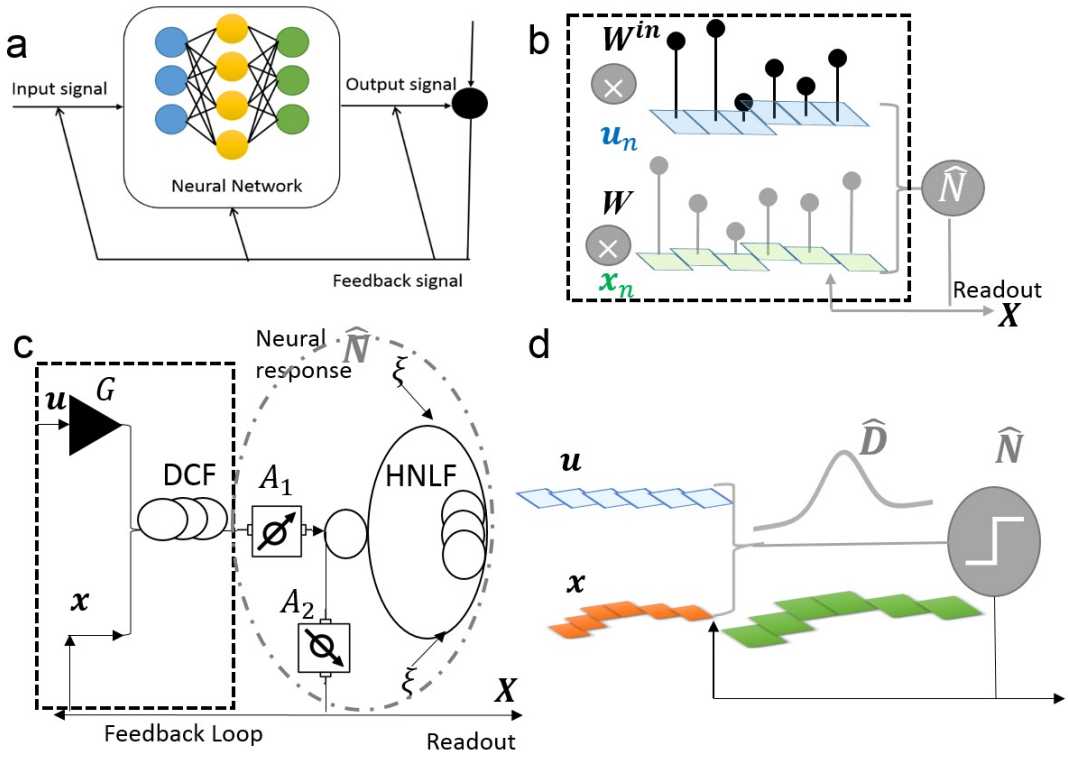

Fig. 1. Design. (a) Neural network concept; (b) Echo state network (ESN), where the main parts of the processing signal mixing and nonlinear transformation are denoted by dashed square and grey circle correspondingly; (c) fiber-optic ESN analogue (FESNA) for dual-quadrature signal processing with parameters: modulated pump ($\xi = 1 + 1i$), highly nonlinear fiber (HNLF) parameters $|\xi|^2/2\gamma L = 2.2\pi$, attenuation coefficients ($A_1 = -18.5dB, A_2 = -6dB$), gain $G = 24dB$, and dispersion compensating fiber (DCF) inducing signal dispersion of $D = 10ps^2$; (d) impact of the proposed scheme on the signal: upon combining the input signal u with the accumulated feedback x, both undergo dispersion induced pulse broadening $\hat{\mathbf{D}}$, which results in the effective mixing of signals with the acquired weights distributed in sin/cos shapes depending on the distance in time between the samples. This represents the weight matrices $\mathbf{W}$ and $\mathbf{W}_{in}$ (the difference between which is controlled by the amplifier G applied to the input signal u). After the subsequent nonlinear transformation the resulted signal is fed back becoming updated x and also collected at the receiver.

higher bandwidth compared to previous realisations (e.g. based on semiconductor laser) and for the first time enabled dual-quadrature signal processing.

This was achieved by utilizing nonlinear loop mirror configuration [7], which is effectively used for optical signal processing [8], [9], [10]. By choosing pump powers and fiber length to enable 2π-phase shift, a nonlinear sine-transformation of both signal quadratures in two branches of the fiber interferometer was achieved.

However in [3], the weights ($\mathbf{W}$ and $\mathbf{W}^{in}$) were assumed to be applied electronically, similar to [2] or [11], [12], [13]). Thus, although fiber enables processing of high bandwidth signals, previous setup was limited by the bandwidth of the available electrical components for signal masking. Here we present a solution for this problem (see Fig. 1(c) dashed square) by utilizing fiber dispersion to achieve synaptic weights. This enables an all-optical FESNA realisation.

To realize signal masking, the injected signal is amplified with amplification parameter G and combined with the feedback signal x, whereupon the resulted signal $v = Gu + x$ goes through dispersion compensating fiber (DCF) with the fixed dispersion coefficient:

$$\hat{\mathbf{D}}[v(t)] = \sqrt{\frac{i}{2\pi D}} \int_{-\infty}^{\infty} v(t') \exp\Big(-i\frac{(t-t')^2}{2D} \Big) \quad (1)$$

Thus, for each symbol duration t_s an effective signal mixing covers Dt_s^{-2} symbols. Depending on the sampling rate at the receiver N_s, there are $N = N_s D t_s^{-2}$ mixed samples, which can be viewed as an analogue of the reservoir size or the number of virtual neurons.

As a result of mixing, and as schematically depicted in Fig. 1d), the dispersion-mixed signal acquires more variation between samples than the injected u. The dispersion-induced masking represents the weight matrices $\mathbf{W}$ applied to real and imaginary components with sine/cosine-based variation enabling mixing between the signal components. The variation between the weight matrices $\mathbf{W}$ and $\mathbf{W}_{in}$ is achieved through the change in the amplifier gain G and the dispersion D. The higher/lower value of D results in a smoother/sharper signal mixing and larger/smaller memory. The purpose of the amplifier G is to induce different mixing between the input and feedback signals. Without the amplifier, using dispersion only, we receive what we call dispersion-managed FESNA (DM-FESNA) and although its performance is similar to the ideal FESNA (with optimal electronic masking) at high signal powers, it decreases quickly at smaller signal powers. The insertion of the amplifier and optimization of its gain parameter G, further referred as optimized DM-FESNA (ODM-FESNA), enables to achieve as good performance as the ideal FESNA in the operating regime of interest (around $BER = 10^{-3}$).

Whereupon mixing, signal undergoes the nonlinear transformation, where both transformed quadratures experience sine-based transformation of real/imaginery parts in upper/lower branches of the fiber-based interferometer:

$$\hat{\mathbf{N}}[v] = \sqrt{A_2/2}(\widehat{\mathbf{HNLF}}[v^u] + i\widehat{\mathbf{HNLF}}[v^l])$$

where $\widehat{\mathbf{HNLF}}[v]$ denotes the Kerr-nonlinearity of the HNLF.

The resulted output is fed back into the system as a new x, thus, resulting in the iterative process of updating ESN states and inducing memory. Also a copy is collected at the receiver. Note, that ESN does not require any additional feedback apart from the transformed signal x. This brings an important advantage of the ESN implementation as it requires only one feedback loop, which is formed directly from the transformed signal.

At the receiver the processed signal is sampled collecting a sequence of samples for each symbol and thus, creating the augmented matrix $\mathbf{X}$. Linear regression is performed electronically to extract symbols from augmented matrix $\mathbf{X}$. Unlike conventional neural networks, which require complex training and processing at the receiver, the advantage of ESN architecture is in the reduced complexity training, which requires only an optimization of the output weights ($\mathbf{W_{out}}$) through a simple linear regression or matrix inversion, namely for a given target function Y_t:

$$\mathbf{W_{out}} = \mathbf{Y_t}\mathbf{X}^T(\mathbf{X}\mathbf{X}^T + \lambda\mathbf{I})^{-1}$$

where λ is a regularization factor and $\mathbf{I}$ is an identity (or unit) matrix.

The sampling rate at the receiver is an important optimization factor, which here was chosen to be $N_s = 32$ samples per symbol (SpS). This is typical for optical communication systems, where $N_s = 2$ SpS is sufficient for the simplest linear equalization (LE), while more advanced signal processing techniques, such as Digital Back Propagation (DBP), usually requires $N_s = 16$ SpS. The parameters of DCF and HNLF fibers, as well as, pumps are optimisation parameters, the values of which will depend on the signal properties and type of distortions. Here for a case of distortions compensation in optical communication links, we chose the standard off-shelf available highly nonlinear fiber, with nonlinear, dispersion, and attenuation coefficients of 10.5 1/W/km, -1.5 ps/nm/km, and 0.8 dB/km, respectively. 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 $P_\xi \gamma_{HNLF} L_{HNLF} = 2.2\pi$ (here we used pump power $P_\xi = 1W$ and a fiber length of $L_{HNLF} = 650m$)). Overall, the effect of dispersion was negligible for the considered signal properties. Note, that higher values of nonlinear coefficients are also available, including some twice as high, which will further reduce the HNLF length.

Table 1: DM-FESNA system parameters

Samples per symbol	32
Maximum phase change	0.01
Pump power	1 W
HNLF Length	650 m
HNLF Attenuation coefficient	0.8 dB/km
HNLF CD	1.5 ps/nm/km
HNLF Nonlinearity	10.5 1/W/km
Attenuators' parameters	$A_{1,2} = -18.5, -6dB$
Amplifier gain	$G = 24dB$
DCF Dispersion	$D = 10ps^2$

III. RESULTS

We demonstrate an application area of the proposed all-optical FESNA technology by using it for distortion compensations. We start with a typical fiber-optic communication system parameters: a single 30 GHz channel modulated with root-raised cosine pulses having 0.1 roll-off transmitted over 100 km single span with varied signal power, see Fig. 2(a). In Fig. 2(b) we study the spectrum broadening due to transformations: one can see that the broadening is small. Here we focus on single channel operation, leaving multichannel regime to further research. The parameters are chosen to demonstrate processing efficiency of the the proposed technology for the record-high bandwidth. The optical fiber link of communication system has 17 ps/nm/km dispersion, 0.2 dB/km attenuation and 1.4 1/W/km nonlinear coefficients. Our focus here is the nonlinear transmission regimes, so we use high input signal powers. We use 1000 symbols for training and 2^{21} QAM-modulated symbols for testing. The optimum number of symbols for training was found to be 1000 symbols and was determined by simulations (as smaller numbers resulted in poorer performance, while larger numbers have led to overfitting). The number of symbols for training was chosen to give at least 100 errors for each error rate calculation point. To quantify the performance, we use bit error rate (BER) as a performance metric. We compare FESNA performance with the standard linear equalization (LE), which compensates for the circular phase shift and other linear distortions. Note, LE represents the minimum distortion compensation technique that is typically employed, presenting the lowest performance level compensation.

Table 2: Communications system parameters

Maximum phase change	0.01
RRC Roll off	0.1
Bandwidth	30 GHz; 100 GHz; 1 THz
Length	100 km; 1 km; 0.5 km
Attenuation coefficient	0.2 dB/km
CD	17 ps/nm/km
Nonlinearity	1.4 1/W/km

The performance comparison is plotted in Fig. 2(a), which shows that the gain in BER is growing for higher nonlinearity and higher order formats. Here we study the ideal FESNA (electronic ESN algorithm with fiber nonlinearity as a neuron activation function) compared to an optical dispersion-managed DM-FESNA and optimized ODM-FESNA. One can

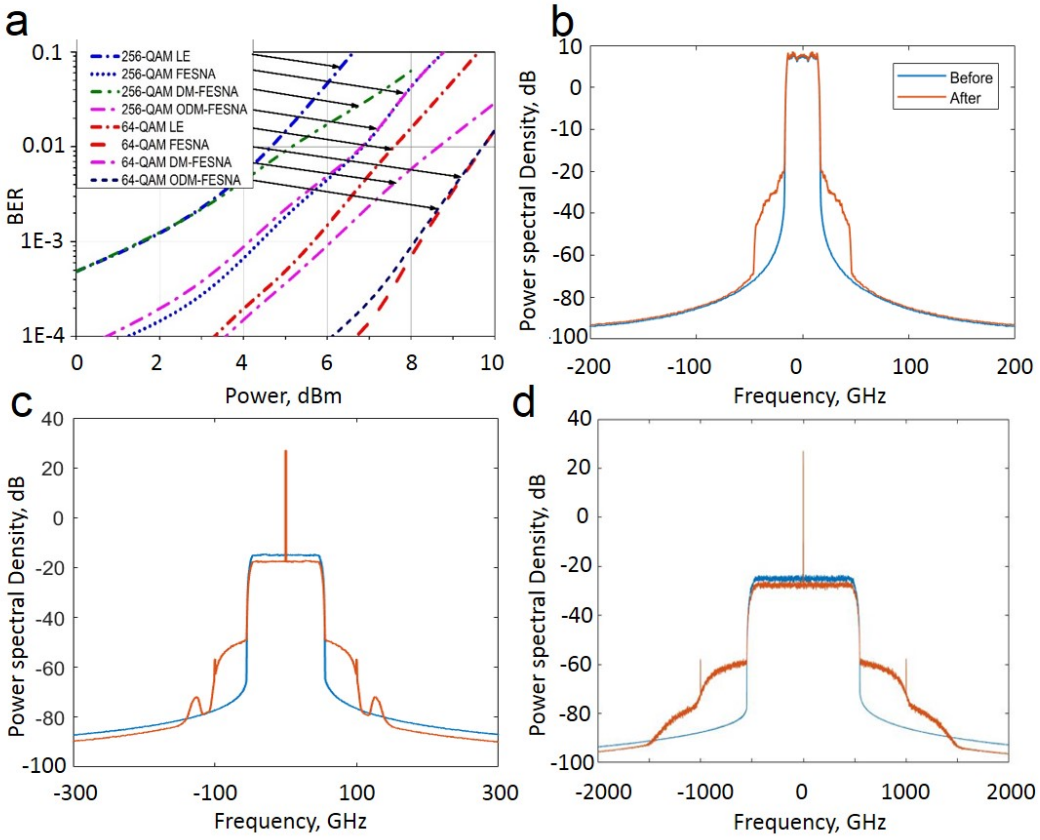

Fig. 2. (a) Performance comparison of different distortion compensation techniques: Linear Equalisation (LE) and ideal Fiber Echo State Network Analogue (FESNA) with its optical realisations: Dispersion-Managed FESNA (DM-FESNA) and optimized DM-FESNA (ODM-FESNA) for 64- and 256-QAM signal transmitted through 100 km of standard single-mode fiber; (b) Spectra of 30 GHz signal before/after (blue/red) ODM-FESNA; (c) Spectra of 100 GHz signal before/after (blue/red) ODM-FESNA and (d) Spectra of 1 THz signal before/after (blue/red) ODM-FESNA.

see that for very high nonlinearity the performance of DM-FESNA is similar to that of ideal FESNA, but starts to decrease to that of LE-performance for lower powers. Thus, if one is interested in performance of BER of the level of 10^{-3} and lower, one needs to use ODM-FESNA, performance of which is comparable to that of the ideal FESNA. Comparing performance for different formats, we see that the higher the order of QAM, the better the gain. Indeed one can see that at BER level of 10^{-3} when compared to LE only, FESNA offers 2 dB and 3dB gain for 64- and 256-QAM correspondingly. This is because higher QAM is associated with higher nonlinearity (more sensitive to signal-to-noise ratio) and this gives FESNA more operational gain. This is a very important feature as for the same complexity FESNA can give a higher gain for higher order formats making it highly relevant for future optical systems, which move to more complex formats [14]. Given that machine learning algorithms and, in particular, neural networks are being developed for optical communications [15], [16], the proposed optical FESNA can be a promising technology for this area. As an additional demonstration of the performance, we plot a spectrum of 256-QAM 30 GHz format in Fig. 2(b) and constellation symbols for 64- and 256-QAM in Fig. 3(a,b) with LE only (blue) and ODM-FESNA (red).

Furthermore, we demonstrate the applicability of the setup to process higher bandwidth signals, which could be relevant for future communication systems, where ultra-wide band communications and high-bandwidth receivers are in development [14]. Moreover, to illustrate the ability of the proposed technology to process high bandwidth signals further we include the study of 100 GHz and 1 THz signals. This demonstration is a necessary step for future applications of neurmorphic technologies in optical communications. Although, here we are confided to a single-channel operation, the proposed technology DM-FESNA is at the core of multichannel signal processing [5]. As here we consider a single-channel operation, we reduce the transmission distance to avoid the effect of higher order nonlinear effects (such as Brillouin and Raman scattering). Nevertheless, the considered set of parameters (including higher order QAM) enables to demonstrate the ability to mitigate nonlinear distortions. Here we use similar transmission setup increasing signal bandwidth to 100 GHz bandwidth, see Fig. 2(c), with the same modulation format, see Fig. 3(c), and 1 THz bandwidth in Fig. 2(d) with 1024-QAM modulation format in Fig. 3(d) with varied transmission distance 1 km and 0.5 km for 100 GHz and 1THz signals correspondingly with the increased input power of 20 dBm. At these parameters BER has improved from 0.027 and 0.012 using LE only to the level of 10^{-3} with the proposed ODM-

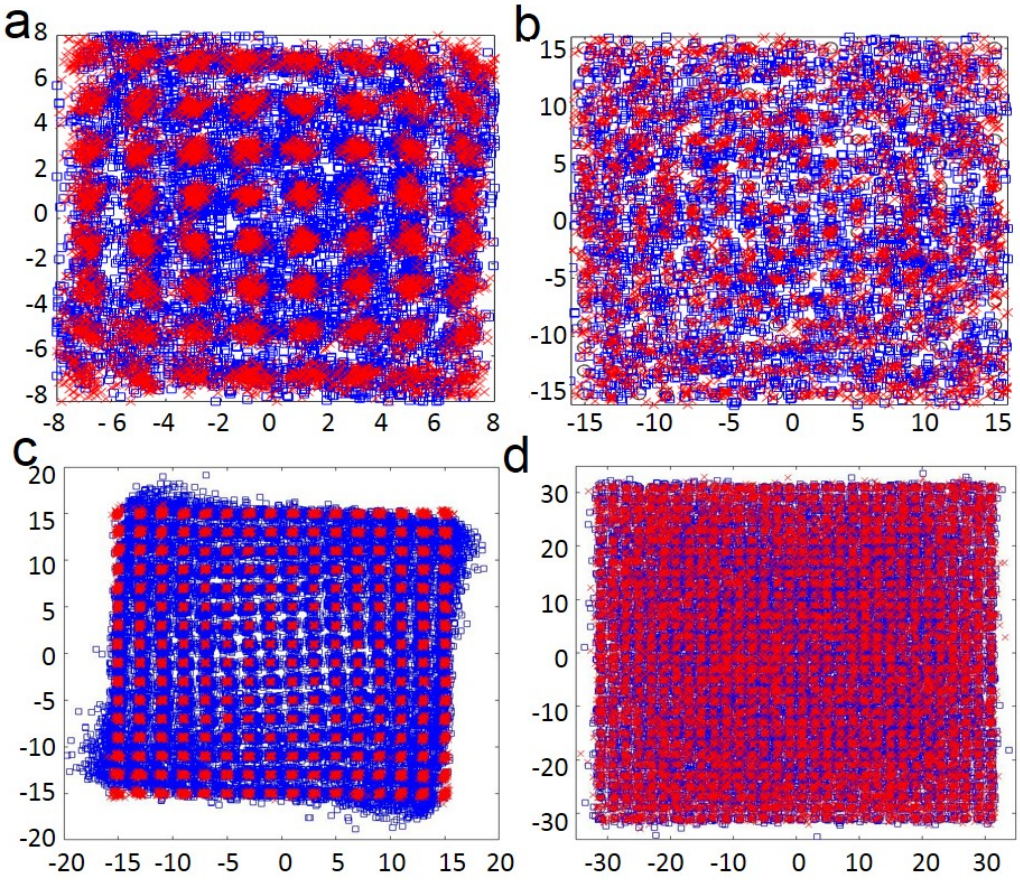

Fig. 3. (a) 64- and (b) 256-QAM 30 GBaud signals transmitted through 100 km fiber at 10 dBm input power after LE (blue) and ODM-FESNA (red); (c) 256-QAM signals transmitted through 1 km fiber at 20 dBm with LE/ODM-FESNA (blue/red); and (d) 1024-QAM signals with transmission length 0.5 km at 20 dBm input power with LE/ODM-FESNA (blue/red).

FESNA processing. Here we used the fixed ODM-FESNA setup for various input signal parameters. The aforementioned examples, demonstrate the efficiency and flexibility of the system when processing high bandwidth signals.

One can see that the proposed FESNA enables significant improvement in performance and is modulation format transparent, which is of high importance for modern communication systems. Overall, the figure illustrates that all-optical neural network realization based on FESNA enables high-efficiency mitigation of linear and nonlinear distortions for high-bandwidth signals.

IV. Conclusions

The proposed design is the first all-optical fiber-based neuromorphic technology for high bandwidth processing (reaching 1 THz). We demonstrate its application for distortion mitigation in optical communications with over 2 dB gain for 64-QAM and 3 dB gain for 256-QAM formats. The technology enables to perform neural network based signal processing in optical domain. It is scalable to different power levels and modulation formats.

Mariia Sorokina is acknowledging support from the Royal Academy of Engineering Research Fellowship RF\201718\17154.

References

[1] H. Jaeger, "Adaptive Nonlinear System Identification with Echo State Networks," Advances in Neural Information Processing Systems **15**, 593-600 (2003).

[2] A. Argyris, J. Bueno, I. Fischer, "Photonic Machine Learning Implementation for Signal Recovery in Optical Communications," Scientific reports **8**, 8487 (2018).

[3] M. Sorokina, S. Sergeyev, and S. Turitsyn, "Fiber Echo State Network Analogue for High-Bandwidth Dual-Quadrature Signal Processing," Opt. Express **27**, 2387-2395 (2019).

[4] M. Sorokina, "High bandwidth all-optical fiber-based neuromorphic signal processing," European Conference on Optical Communications P39 (2019).

[5] M.Sorokina, "Multi-channel optical neuromorphic processor for frequency-multiplexed signals," preprint Researchgate, February 2020 DOI: 10.13140/RG.2.2.33164.46720.

[6] 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).

[7] N. J. Doran and David Wood, "Nonlinear-Optical Loop Mirror," Opt. Lett. **13**, 56-58 (1988).

[8] M. Sorokina, "Design of multilevel amplitude regenerative system," Opt. Lett. **39**, 2499-2502 (2014)

[9] M. S. Sorokina and S. K. Turitsyn, "Regeneration Limit of Classical Shannon Capacity," Nat. Commun. **5**, 3861 (2014).

[10] M. Sorokina, S. Sygletos, and S. Turitsyn, "Regenerative Fourier Transformation for Dual-Quadrature Regeneration of Multilevel Rectangular QAM," Opt. Lett. **40**, 3117-3120 (2015).

[11] F. Duport, A. Smerieri, A. Akrout, M. Haelterman, and S. Massar, "Fully analogue photonic reservoir computer," Scientific Reports **6**, 22381 (2016).

[12] Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman and S. Massar, "Optoelectronic reservoir computing," Sci. Rep. **2**, 287 (2012).

[13] 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).

[14] 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).

[15] 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 n CO-OFDM by ANN-Based Nonlinear Equalization," Opt. Lett. **40**, 5113-5116 (2015).

[16] M. Sorokina, S. Sygletos, and S. Turitsyn, "Sparse Identification for Nonlinear Optical Communication Systems: SINO Method," Opt. Express **24**, 30433-30443 (2016).