Unsupervised and supervised machine learning for performance improvement of NFT optical transmission

1st Oleksandr Kotlyar
Aston Institute of Photonic Technologies
Aston University
Birmingham, UK
o.kotlyar1@aston.ac.uk

2nd Maryna Pankratova
Aston Institute of Photonic Technologies
Aston University
Birmingham, UK
m.pankratova@aston.ac.uk

3rd Morteza Kamalian
Aston Institute of Photonic Technologies
Aston University
Birmingham, UK
kamaliam@aston.ac.uk

4rd Anastasiia Vasylchenkova
Aston Institute of Photonic Technologies
Aston University
Birmingham, UK
vasylcha@aston.ac.uk

5th Jaroslaw E. Prilepsky
Aston Institute of Photonic Technologies
Aston University
Birmingham, UK
y.prylepskiy1@aston.ac.uk

6th Sergei K. Turitsyn
Aston Institute of Photonic Technologies
Aston University
Birmingham, UK
s.k.turitsyn@aston.ac.uk

Abstract—We apply both the unsupervised and supervised machine learning (ML) methods, in particular, the k-means clustering and support vector machine (SVM) to improve the performance of the optical communication system based on the nonlinear Fourier transform (NFT). The NFT system employs the continuous NFT spectrum part to carry data up to 1000 km using the 16-QAM OFDM modulation. We classify the performance of the system in terms of BER versus signal power dependence. We show that the NFT system performance can be improved considerably by means of the ML techniques and that the more advanced SVM method typically outperforms the k-means clustering.

Index Terms—Machine learning, support vector machine, k-means clustering, nonlinear Fourer transform, optical communications

I. INTRODUCTION

ML techniques have recently attracted a great deal of attention as effective tools for mitigation of the nonlinearity-induced signal distortions in fibre-optic communications [1]–[3]. ML can be applied to establish high dimensional relationships between system components (parameters) while taking various factors and constraints into account simultaneously [3]. In our study here we address the NFT-based communication systems employing the continuous nonlinear spectrum [4] and append those systems with the ML block at the receiver (Rx) side after the demodulation.

We assume here that the propagation of the optical signal in the fibre is described by the lossless nonlinear Schrodinger

equation (NLSE). We write down the NLSE for the slow-varying envelope function $q(z,t)$ in the dimensionless form (see e.g. Ref. [4] for normalisations):

$$iq_z - q_{tt} - 2|q|^2 q = \eta(z,t), \qquad (1)$$

where z is a normalised distance along the fibre, t is time and $\eta(z,t)$ is the optical noise term arising due to the amplification (the noise is assumed to be distributed along the link). Eq. (1) without the right hand side can be solved using the NFT technique [5], within which the special "nonlinear eigenmodes" propagate inside the nonlinear Fourier (NF) domain in a linear decoupled manner while the evolution in the optical domain occurs according to nonlinear Eq. (1). The linear NFT modes evolution can be utilised in the effectively nonlinearity-free NFT-based transmission framework [4], [6], where we use the parameters of these eigenmodes as information carriers. The nonlinear frequency division multiplexing concept (NFDM) was proposed in [6] which deals with the multiplexing of NF spectrum components, in contrast to the conventional WDM dealing with Fourier modes. The signal components (eigenmodes) inside the NF domain represent the two orthogonal types: the continuous nonlinear waves and non-dispersive solitons. In our current study we use only the former continuous part of the NF spectrum for the modulation, similarly to the systems proposed in [7], [8], i.e. by using the so-called nonlinear inverse synthesis (NIS) method, Fig. 1. Within the NIS method, the signal is generated at the transmitter (Tx) from the modulated and encoded continuous nonlinear spectrum profile $r(\xi)$ by employing the inverse NFT. In the expression for $r(\xi)$ the parameter ξ plays the role of the nonlinear analogue of frequency. Then the generated signal is launched into the optical fibre. At the Rx, located at distance $z = L$, we perform the NFT operation recovering
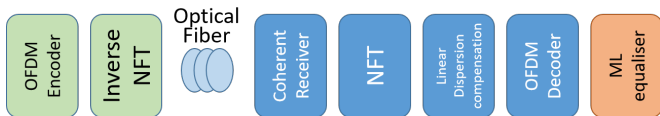
Fig. 1. Schematic of the nonlinear frequency division multiplexing (NIS) transmission scheme appended with the ML equalisation at Rx side.

the spectral profile $r(\xi, L)$, and then remove the accumulated phase rotation:

$$r(\xi, 0) = r(\xi, L)e^{-4i\xi^2 L}.$$

In our work here we use the popular OFDM modulation to construct the $r(\xi)$ profile [7], [8], with the effective bit-rate of the resulting system being 52 Gbps.

The goal of this paper is to study the performance improvement of the OFMD-NIS system described above by using the ML techniques. The results are qualified via the BER vs. power curves and the performance improvement obtained through the use of supervised (SVM) and unsupervised (k-means clustering) ML methods. We note that the SVM method has been successfully applied for the improvement of "conventional" optical systems [9], [10], while the k-means clustering method was also studied in earlier works [11], [12] but is still of interest [13]. However, to the best of our knowledge, application of ML-based equaliser at the Rx side to NFT-based transmission (in the NF domain) has not been reported yet. Only a time-domain neural network receiver for NFDM system has been proposed [14].

II. SIMULATIONS RESULTS

For our simulations we use an OFDM system with 128 sub-carriers as

$$r(\xi, z = 0) = \sum_{k=-63}^{64} c_k \operatorname{sinc}(2\xi - k), \qquad (2)$$

where c_k are 16-QAM symbols. The transmission takes place in burst mode [8], where the full temporal support of our single symbol (burst) is 9.8 ns. We assume that the transmission occurs down the single mode fibre with standard parameters [4], [6]. The propagation distance of our system is 1000 km. Overall, for the data detection and the comparison of the performance we consider approximately 5000 realizations of inline noise and randomly selected 16-QAM coefficients.

Received symbols were identified using conventional hard decision detection, concretely, the complex plane of received c_k was divided into the equal squares, and the symbols getting into the wrong square were identified as errors. We also use a phase equalisation, meaning the uniform rotation of the whole constellation combined with hard decision method. The BER vs. launch power curves obtained in the NFDM systems without ML is presented in Fig. 2 by the amber (without phase rotation) and blue curves.

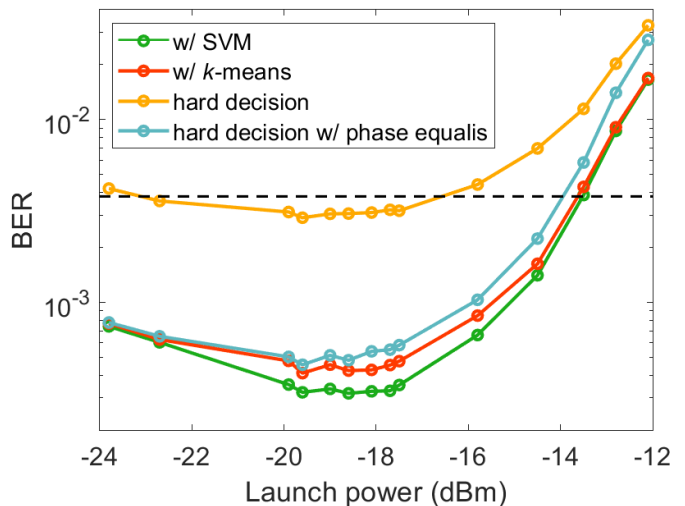It is known that the NFDM systems are impacted by the noise on a larger scale then the conventional systems, the

Fig. 2. BER vs optical launch power for 16-QAM NFDM system for conventional hard decision, hard decision after phase equalisation, k-means and SVM equalisers (amber, blue, red and green curves respectively). Black dashed line represents HD-FEC threshold

fact which severely limits their performance [15]. To improve the performance and noise tolerance of the NFDM systems we apply two ML techniques for data detection inside the NF domain. First we use unsupervised ML algorithm, the k-means clustering that can be applied to unlabelled data i.e., the data that does not belong to any group. The goal of this algorithm is to group data into clusters and give them some labels. The algorithm starts from random cluster centroids (centres) initialisation. Then the points closest to the centroid are assigned to appropriate cluster. After that, the locations of centroids are re-initialised to the points corresponding to the value that will minimize the average of the distances belonged to cluster centroid. We repeat the procedure of re-initialization till centroid's position does not change.

In our paper we use the so-called *k-means++* algorithm [16] for cluster centre initialisation implemented in MATLAB. To search for the lower local minima we perform more than 20 replicates of repeat clustering (typically 25–30 depending on the sum of minimal distances) using the new initial cluster centroid positions to avoid termination at local minima. Then, the clusters configuration corresponding to the minimum of the total sum of distances between centroids and symbols which belong to centroids, was used for the BER calculation. In Fig. 3 we present the resulting cluster configuration for the launch power -17.5 dBm (the clusters are marked with a specific colour), where the points which were incorrectly detected by hard decision method with phase equalization are highlighted by the black circles.

Next we aim to improve the performance of the system by employing the supervised ML algorithm and process our data at the Rx side by using the SVM method. SVM is known as a popular powerful ML tool for classification and regression tasks. One of the main features of SVM is that the algorithm creates decision boundaries with the largest possible
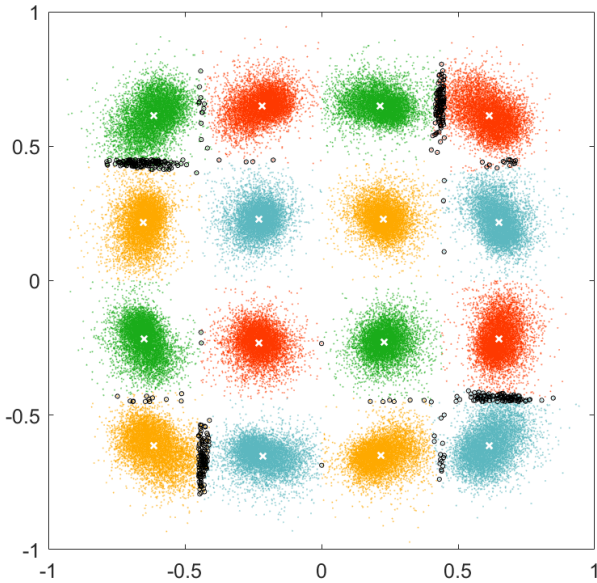
Fig. 3. Cluster obtained by k-means method with centroids (white crosses) and points that were detected incorrectly by hard decision with phase equalisation and correctly by k-means (black circles) for power is equal to -17.5 dBm
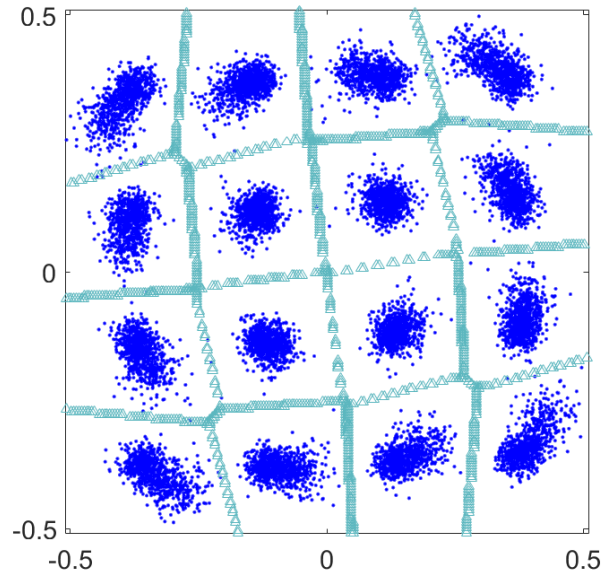

Fig. 4. Resulting nonlinear decision boundaries obtained by SVM method for the power is equal to -18.6 dBm (corresponds to minimum BER)

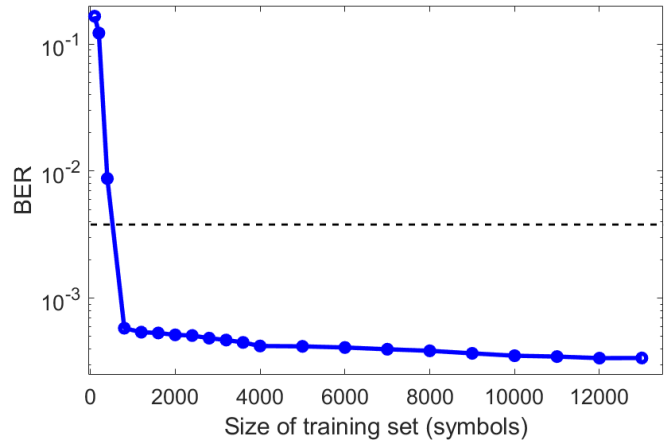Fig. 5. Impact of the number of training symbols on SVM model performance. Black dashed line represents HD-FEC threshold.

margins for the training samples from different clusters. Then the new samples are assigned to one of the groups depending on the encircled domain they fall in. In our paper we use the open source LIBSVM package [17] for MATLAB. The typical implementation of SVM consists in two essential steps. In the first step, we train a data set to obtain a model. Around 10,000 training symbols (that is 8% of the overall number of samples) were used to train the system. In the second step, the model obtained was applied for identification of the rest of the data set. Therefore, the BER of the system was evaluated from the data excluding 8% the points used for the training. The decision boundaries obtained by the SVM method differ significantly from the ones (the straight lines) used in the hard decision method as shown in Fig. 4.

The BER resulting from applying the four methods described above are presented in Fig. 2 for various signal powers at 1000 km.

The k-means can correctly classify some points that are misclassified by the conventional detection, represented by black circles in Fig. 3: the algorithm can recover some additional fraction of data on the edges of clusters. Therefore, we see that the application of k-means clustering provides slightly better performance than conventional detection with the phase shift of the received data points depicted in Fig. 2. The constellation points with the largest relative power (the corners of the constellation) contribute more to the overall error, such that the k-means correction in these areas is most noticeable. Overall, it can be seen from Fig. 2 that SVM delivers the best performance for our system.

The influence of the size of our training set on the BER for the signal power of -17.5 dBm is shown in Fig 5. The overall number of symbols used for SVM simulation in the

training and test set is equal to 131,072. It is clear from that figure that the performance changes insignificantly for the test sets larger than 4000 symbols. Therefore our choice of 10,000 symbols for the model training, which lies in the saturation region in Fig. 5, guarantees the high enough accuracy of the SVM equaliser used for drawing Fig. 2.

III. CONCLUSIONS

We have applied both supervised and unsupervised ML methods as equalisers at the Rx side for the data detection in the NFT-based optical transmission system. Two ML methods, the SVM and k-means clustering, were compared with a hard decision detection scheme. It was shown that the SVM provides the best performance improvement which is demonstrated by means of our comparing the BER vs power dependences for SVM, k-means and conventional detections

schemes. The k-means clustering albeit improving the system performance, results in a generally lower value of improvement compared to the SVM-based equalisation. The overall BER improvement achieved by the SVM-based equalisation was about 35% at the optimal power.

REFERENCES

[1] F. Musumeci, et al., "A Survey on Application of Machine Learning Techniques in Optical Networks," arXiv preprint, arXiv:1803.07976v3, October 2018.

[2] B. Karanov et al., "End-to-End Deep Learning of Optical Fiber Communications," J. Lightwave Technology, vol. 36, pp. 4843–4855, October 2018.

[3] D. Zibar, M. Piels, R. Jones, and C.G. Schäeffer, "Machine learning techniques in optical communication," J. Lightwave Technol, vol. 6, pp. 1442–1452, March 2016.

[4] S. K. Turitsyn et al., "Nonlinear Fourier transform for optical data processing and transmission: advances and perspectives," Optica, vol. 4, pp. 307–322, March 2017.

[5] V. E. Zakharov and A. B. Shabat, "Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media," Sov. Phys.-JETP, vol. 34, pp. 62–69, January 1972.

[6] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear Fourier transform, Part I: Mathematical Tools," IEEE Trans. Inform. Theory, vol. 60, pp. 4312–4328, July 2014.

[7] J. E. Prilepsky, S. A. Derevyanko, K. J. Blow, I. Gabitov, and S. K. Turitsyn, "Nonlinear Inverse Synthesis and Eigenvalue Division Multiplexing in Optical Fiber Channels." Phys. Rev. Lett., vol. 113, pp. 013901-1–013901-5, July 2014.

[8] S. T. Le, I. D. Philips, J. E. Prilepsky, P. Harper, A. D. Ellis, and S. K. Turitsyn, "Demonstration of Nonlinear Inverse Synthesis Transmission over Transoceanic Distances," J. Lightwave Technol. vol. 34, pp. 2459–2466, May 2016.

[9] T. Nguyen, S. Mhatli, E. Giacoumidis, L. Van Compernolle, M. Wuilpart, and P. Mégret, "Fiber Nonlinearity Equalizer Based on Support Vector Classification for Coherent Optical OFDM," IEEE Photonics J., vol. 8, pp. 1–9, April 2016.

[10] J. Thrane, J. Wass, M. Piels, J. C. M. Diniz, R. Jones, and D. Zibar, "Machine Learning Techniques for Optical Performance Monitoring From Directly Detected PDM-QAM Signals," J. Lightwave Technology, vol. 35, pp. 868–875, February 2017.

[11] T. Zhao, A. Nehorai, and B. Porat, "K-means clustering-based data detection and symbol-timing recovery for burst-mode optical receiver," IEEE Transactions on Communications, vol. 54, pp. 1492–1501, August 2006.

[12] N. G. Gonzalez, D. Zibar, X. Yu, and I. T. Monroy, "Optical phase-modulated radio-over-fiber links with k-means algorithm for digital demodulation of 8PSK subcarrier multiplexed signals," 2010 Conference on Optical Fiber Communication (OFC/NFOEC), collocated National Fiber Optic Engineers Conference, San Diego, CA, paper OML3, March 2010.

[13] J. Zhang, W. Chen, M. Gao, B. Chen, and G. Shen, "Novel Low-Complexity Fully-Blind Density-Centroid – Tracking Equalizer for 64-QAM Coherent Optical Communication Systems," Optical Fiber Communication Conference, OSA Technical Digest, paper M1G.4, March 2018.

[14] R. T. Jones, S. Gaiarin, M. P. Yankov, and D. Zibar, "Time-Domain Neural Network Receiver for Nonlinear Frequency Division Multiplexed Systems," IEEE Photonics Technology Letters, vol. 30, pp. 1079–1082, June, 2018.

[15] V. Aref, S.T. Le, and H. Buelow, "Modulation over Nonlinear Fourier Spectrum: Continuous and Discrete Spectrum", J. Lightwave Technol., vol. 36 , pp. 1289 - 1295, March 2018.

[16] A. David, and S. Vassilvitskii. "K-means++: The Advantages of Careful Seeding." SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 1027–1035, January 2007.

[17] Chang Chih-Chung, and Lin Chih-Jen. "LIBSVM: A library for support vector machines." ACM Transactions on Intelligent Systems and Technol., vol. 2, pp. 27:1–27:27, April 2011.