Fibre Optic Nonlinear Technologies [FONTE] - A European Industrial Doctorate [GA766115]

Document Details

Title	Deliverable 5.1 Transmission regime definition and plan of experiments
Deliverable number	D5.1
Deliverable Type	Report (public)
Deliverable title	Transmission regime definition and plan of experiments
Work Package	WP5 – Experimental implementation and testing of NFT systems
Description	Details of the plans for the experimental transmissions of NFT-based systems are presented
Deliverable due date	31 Jul 2019
Actual date of submission	24/12/2019
Lead beneficiary	Aston University
Version number	V1.0
Status	FINAL

Dissemination level

PU	Public	X
CO	Confidential, only for members of the consortium (including Commission Services	

Project Details

Grant Agreement	766115
Project Acronym	FONTE
Project Title	Fibre Optic Nonlinear Technologies
Call Identifier	H2020-MSCA-ITN-2017
Project Website	fonte.astonphotonics.uk
Start of the Project	1 June 2018
Project Duration	48 months

Consortium

EC Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 766115

Executive Summary

In this deliverable, we define the practical challenges of interest that the ESRs are (and have been) planned to tackle in the scope of the project. Since ESRs follows their individual research direction according to the work packages, we defined different experimental plans for each ESR. In the current report, the experimental plans are detailed based on the progress of ESRs in their research. We present detailed experimental plans for ESR 2 (Vinod) and ESR 3 (Stenio) as they had so far a successful progress in their research. Due to the late hiring, the other two ESRs, ESR 1 (Vlad) and ESR 4 (Abtin), are still in early research stages. We defined their experimental plans as well, but more details will be given in the next deliverables.

ESR 1 is investigating on the machine learning (ML)-based approaches to improve the performance of nonlinearity mitigation (NLM) techniques in fibre-optic communication systems, in particular the performance of nonlinear frequency-division multiplexing (NFDM) systems where the information in encoded into the nonlinear Fourier spectrum. In this report, we overview the practical challenges and define the next experimental plans to address these challenges.

ESR 2 follows his research on the impairments mitigation of NFDM systems. NFDM systems are severely suffered from some practical impairments such as non-ideal fiber amplification, and non-ideal transceiver's characteristics. ESR 3 has been working on some solutions to mitigate such impairments. we briefly review the challenges in designing NFDM experiments and define our plans for compensation of the nonlinear distortion arising from transceivers.

ESR 3 is investigating on Reservoir Computing (RC) as a new tool for implementing signal processing in the optical domain. It follows the architecture of a recurrent neural network (RNN), which allows the process of time-dependent signals. The advantage of RC over RNNs is that the recurrent connections are determined randomly (called reservoir) and do not require any changes/training over time. This characteristic makes RC an excellent option to be used in photonic systems. In this report, we define our experimental plans to boost the performance of some RC structure in practice.

ESR 4 is working on machine-learning solutions for optical fiber communications, including NFDM systems. The research direction is to use neural networks and kernel methods for mitigation (equalization) of the impairments of optical fiber. The plan is first to develop such solutions in simulation environment for nonlinear Schrodinger's (NLS) equation, the basic model for an optical fiber. Then, we verify the developed solutions on the real fiber and assess their performance on a high-speed transmission experiment.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ACRONYMS

AIPT	Aston Institute of Photonic Technologies
ASE	Amplified Spontaneous Emission
CD	Chromatic Dispersion
DSP	Digital Signal Processing
EC	European Commission
EID	European Industrial Doctorates
ESR	Early Stage Researcher
FONTE	Fibre Optic Nonlinear Technologies
LPA	Lossless Path-Averaged
ML	Machine Learning
NFDM	Nonlinear Frequency-Division Multiplexing
NLSE	Nonlinear Schrödinger Equation
NFS	Nonlinear Fourier Spectrum
NFT	Nonlinear Fourier Transform
NLM	Nonlinearity Mitigation
OA	Optical Amplifiers
PMD	Polarization-Mode Dispersion
RC	Reservoir Computing
RNN	Recurrent Neural Network
AWG	Arrayed Waveguide Grating
OOK	On-Off Keying
PAM	Pulse Amplitude Modulation
SNR	Signal-to-Noise Ratio
DNN	Deep Neural Network
ANN	Artificial Neural Network

1 ESR 1

The optical fibre channel is a dispersive-nonlinear channel [1]. The developments in the coherent detection hardware and the digital signal processing (DSP) algorithms enabled the effective mitigation of the linear transmission impairments, notably, chromatic dispersion (CD) and polarization-mode dispersion (PMD) [2]. Therefore, the fibre nonlinear response, especially, Kerr nonlinearity, and the amplified spontaneous emission (ASE) noise generated by optical amplifiers (OA) are currently the bottlenecks of the performance of fibre-optic communication systems [3, 4]. In more details, Kerr nonlinearity distorts the signal waveform proportionally to its local intensity, by thus effectively imposing the upper limit on the signal power and, hence, on signal-to-noise ratio (SNR). Consequently, the mitigation of the nonlinear distortion improves the system performance by unlocking higher SNRs. Several nonlinearity mitigation (NLM) methods have been proposed up to date [5-7], including some recent approaches based on machine learning (ML) [8–11]. The conventional NLM methods have a general problem - they don't take nonlinearity into account during system design but consider it as a perturbation needed to be suppressed by the NLM methods.

Another view on the nonlinearity mitigation is taken by nonlinear Fourier transform (NFT)-based approaches [12–18]. In this paradigm, Kerr nonlinearity is considered as the essential property of the channel, not as a destructive phenomenon. The lossless nonlinear Schrödinger equation (NLSE) which governs the propagation of an optical pulse in an ideal optical fiber can be exactly solved using the NFT [19]. The nonlinear spectral components which are obtained by NFT of a signal do not interfere during propagation through a lossless fiber [13,19]. Further, the NFT enables mitigation of the transmission impairments by simple equalization. These capabilities of the NFT to simplify the fiber-optic communication have encouraged new type of transmission techniques that are based on NFT. Nonlinear frequency division multiplexing (NFDM) techniques are transmission techniques that encode data in the nonlinear spectrum [13,15]. As the NFT requires lossless fiber model, NFDM techniques are challenged due to the presence of loss in real fibers (and accordingly ASE noise of amplifiers) [20]. Reliable ways of countering these distortions in NFDM systems are of high demand. Lossless path-averaged (LPA) model [20], the optimization of modulation format [21], and the application of advanced decoding schemes [22] have been proposed to deal with the impact of losses and ASE noise.

1.1 Research and Experimental Plan

The application of ML approaches was recently found especially beneficial to solve the mentioned issues [23–27]. Nonetheless, not so many works have attended the usage of ML in the NFT context. Within this project, we are going to utilise the expertise of the previous approaches, when ML and NFT were employed together, and to extend this paradigm to more capacity-efficient systems. We plan to further tune the system parameters by means of advanced optimisation techniques, including machine learning approaches. Moreover, we are going to develop the novel decoding technique based on ML. Particularly, we will consider both numerically and experimentally the long-haul NFDM system, similar to the one considered in [17]. The considered system (Figure 1) will use both discrete and continuous NFS as the information carriers. Noteworthy, the system [17] did not utilize any ML-based elements. Conversely, we will replace some elements of the aforementioned scheme with the ML units, which are supposed to combine the reduced complexity with the improved performance metrics. Besides, the other ML-based nonlinearity mitigation methods will be considered.

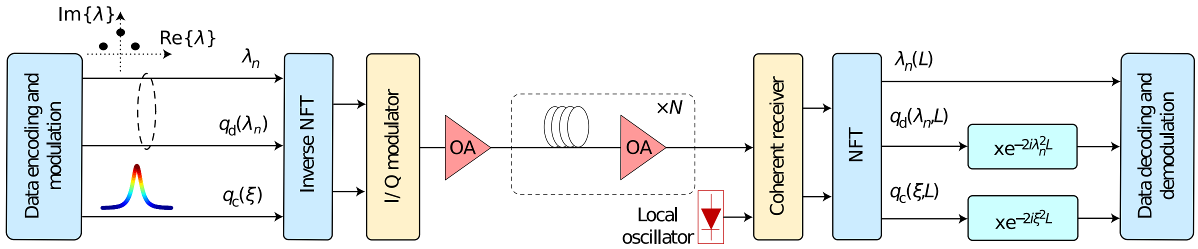

Figure 1: The NFDM system under consideration. The figure is taken from [17].

2 ESR 2

Nonlinear Fourier Transform (NFT) is developed based on nonlinear Schrödinger equation (NLSE) which models propagation of an optical pulse in a lossless optical fiber. Since the practical fiber are lossy, the LPA model is applied in order to make the fiber model suitable for NFT and NFDM systems. In LPA model the lossy propagation model is approximated to a lossless propagation (in a fiber with different nonlinear parameter) [20]. Many NFDM systems were demonstrated in the past by using LPA model. The error associated with the path-average approximation gets stronger at high power, bandwidth and longer spans. Recently, it was shown in simulations that the path-average model can be avoided by designing the NFDM systems using dispersion decreasing fibers (DDFs) [28]. The simulation results present potential advantages of using DDFs in NFDM systems. In this report, first, we briefly review the proposed approach to address the challenge in NFDM systems due to the fiber-loss. Then, we describe some of the challenges in the designing the experiments and possible approach to address them.

2.1 NFDM System Using DDF

It is well-known that the lossless fiber-model belongs to the class of integrable nonlinear evolution equations and can be solved exactly using NFTs [19]. In the presence of fiber-loss, which is always the case in real fibers, the fiber-model is not exactly solvable using NFTs [20]. It was shown that if a fiber with varying parameters is designed then for certain conditions the fiber-model will become integrable [29]. Thus, the fiber-model can be exactly solved using NFT. The propagation of a pulse $Q(\ell,t)$ in such a fiber can be modelled using the NLSE with varying dispersion and nonlinear parameters [29]

$$\frac{\partial Q}{\partial \ell} + i\frac{\beta_2(0)D(\ell)}{2}\frac{\partial^2 Q}{\partial t^2} - i\gamma(0)R(\ell)|Q|^2 Q = -\frac{\alpha}{2}Q. \tag{1}$$

where ℓ represents the propagation distance and t is retarded time. The parameters α, β_2 and γ are the loss, dispersion and nonlinear parameters respectively. $\beta_2(0)$ and $\gamma(0)$ are the initial values (i.e. at $\ell = 0$) of dispersion and nonlinear parameters. $D(\ell)$ and $R(\ell)$ are the normalized profiles that introduces length dependent variation in the dispersion and nonlinear parameters. Here, we consider the case of anomalous dispersion (i.e. $\beta_2(0)D(\ell) < 0$). For the case of fiber-links with lumped amplification, above condition can be satisfied if the following relation holds

$$\frac{\beta_2(\ell)}{\gamma(\ell)} = \mathrm{A}\, e^{-\alpha\ell}, \tag{2}$$

Where $\beta_2(\ell) = \beta_2(0)D(\ell)$, $\gamma(0) = \gamma(0)R(\ell)$ and $A = \beta_2(0)/\gamma(0)$.

Such fibers with varying parameters were designed earlier for classical-soliton based systems [30,31]. The profile of such a fiber is shown in Figure 2, for $\alpha = 0.2$ dB/km and for realistic values of fiber radius and dispersion and nonlinear parameters at $\ell = 0$.

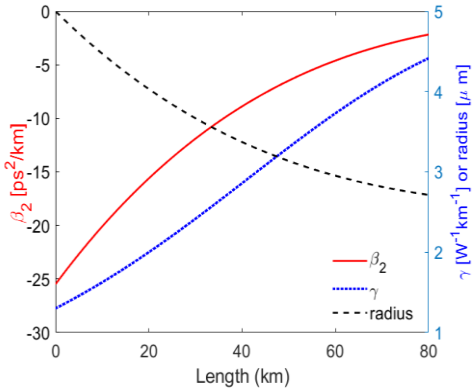

Figure 2: Parameters of dispersion decreasing fiber of 80 km length.

An NFDM system designed using dispersion decreasing fiber avoids the approximation error due to LPA model, hence, have performance advantage which was shown in simulations. In order to validate the simulation results using experiments, we need fibers for which condition given in [28,29] is satisfied.

2.2 Research and Experimental Plans

In order to see the actual advantage of using DDFs, the transmission experiments need to be planned properly. The gains could be visible in experiments if the NFDM system is designed with longer span. However, the DDFs currently available in-house are of very short-lengths (around 1 km). Furthermore, it is crucial for NFDM systems that the signal generated at transmitter should be free from distortions. In experimental systems, the signal generated at transmitter suffers from distortions that should also be minimized. The distortions at transmitter comes from various factors like limited bandwidth and nonlinear response of components. There are various methods to pre-distort waveform digitally in order to suppress the undesired response of the overall transmitter [32-36]. Volterra series are well known for modelling nonlinear systems with memory. It can be used to model the cumulative effect from the transmitter devices. One of such technique is indirect learning architecture (ILA) which is shown in Figure 3 [36].

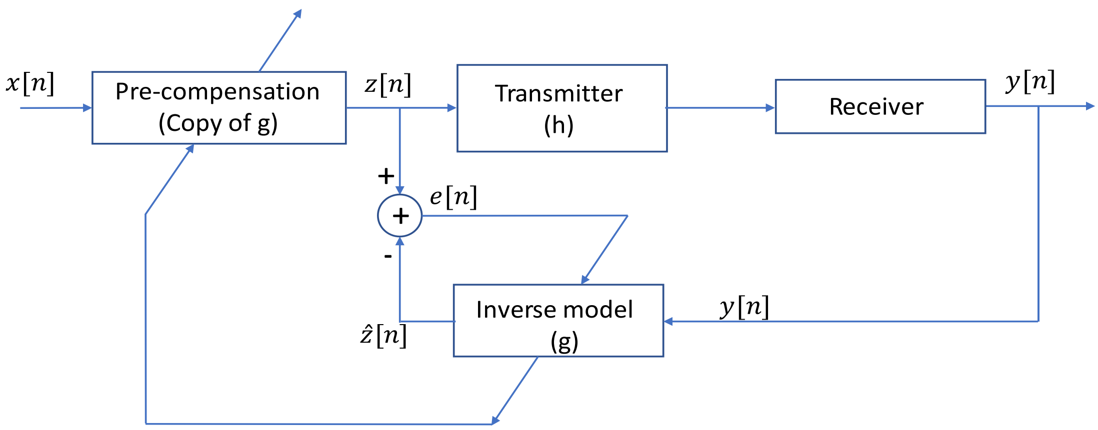

Figure 3: Schematic of indirect learning architecture.

In this approach, the inverse model of transmitter is obtained iteratively using some optimization algorithm such as recursive lease squares. A pre-compensated sequence $z[n]$ is obtained from the input sequence x[n] by applying the pre-compensation Volterra filter g

$$z[n] = g * x = \sum_{p=1}^{P} \sum_{m=-\frac{Mp-1}{2}}^{\frac{Mp+1}{2}} g_{m,p} x_i^p [n-m] \tag{3}$$

where P denotes the nonlinearity order, M_p denotes the memory of the filter at pth order and $g_{m,p}$ is the Volterra kernel coefficient of the corresponding memory tap m and nonlinear order p. The signal at the output of system is $y[n]$ which is impaired with transmitter distortions (assuming receiver is ideal)

$$y\,[n] = h * z[n].$$

These output signal samples $y[n]$ are used for training the pre-compensation filter g. The training model is updated iterative in order to minimize the error

$$e[n] = z[n] - \hat{z}[n]\,.$$

The error $e[n]$ approaches to zero when the inverse of transmitter model is obtained i.e. $g = h^{-1}$. Once the error is minimized, a pre-compensation filter is obtained which compensates the transmitter distortion.

At Nokia Bell labs, we propose the following experiment plans:

1. Earlier, we have shown performance advantage in an NFDM system designed dispersion decreasing fiber in simulations [28]. We want to perform experiments with it but the in-house available fiber does not fit with the requirement. This experiment plan depends upon the availability of the required fiber.
2. We plan to work on the compensation of transmitter nonlinearity by building on the existing approaches. We focus on achieving a greater performance advantage in overall end-to-end system by tuning the existing approaches as well as by using novel optimization algorithms.
3. Further, we will explore the use of machine learning techniques such as deep neural network in order to obtain performance advantage with reduce complexity in comparison to the conventional methods.

3 ESR 3

Reservoir computing (RC) is a new tool for implementing signal processing in the optical domain. It follows the architecture of a recurrent neural network (RNN), which allows the process of time-dependent signals. The advantage of RC over RNNs is that the recurrent connections are determined randomly (called reservoir) and do not require any changes/training over time. This characteristic makes RC an excellent option to be used in photonic systems. However, the relation between the RC and optical fiber impairments is not clear. To get a better grasp of its functionality, we have started the investigation with a digital implementation of RC aiming at signal equalization for direct-detection intensity-modulated optical transmission systems. To further improve results, we also proposed the use of RC together with an optical structure to pre-process the signal. This optoelectronic implementation uses an arrayed waveguide grating (AWG) to split the signal and greatly improved the performance of the system, as proved by simulations.

The RC is divided into two components. The first part is the reservoir, an RNN, which is initialized with random weights and interconnections and kept fixed. Its function is to transform the input signal from one state to another, through nonlinear dynamics. The second is the weights that connect the reservoir's output with the readout functions. It is the only part that is trained and transforms the reservoir's output into the desired signal.

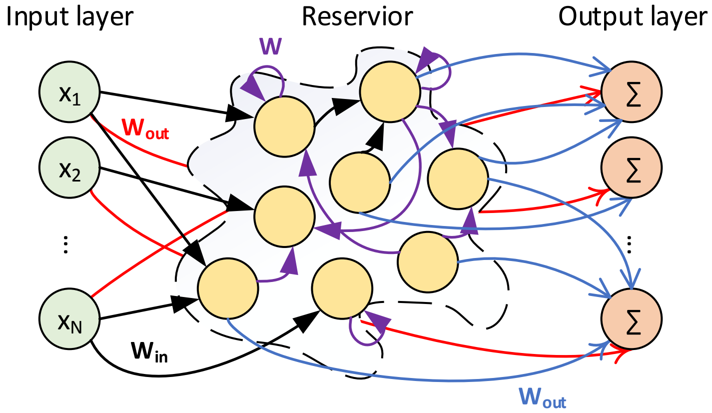

Figure 4: Representation of a reservoir topology.

Figure 4 illustrates the structure of reservoir computing used in the numerical results. The input signal goes into the reservoir with a linear transformation given by W_{in} (defined randomly). A bias is typically added to make the signal operate in different regions of the nonlinear function (the bias' weights are considered in W_{in}). The square matrix W defines the number of neurons and how they are connected. The output layer has two different sets of weights — first, the connection between the input and the readout functions. Second, the connection between reservoir and readout functions. Both weights are concatenated and described by W_{out}. Equation 1 describes the space-state equations for this structure.

$$x[n] = \tanh(W_{in} \cdot u[n] + W \cdot x[n-1]) \tag{4}$$

$$y[n] = W_{out} \cdot [u[n]; x[n]]$$

3.1 Research and Experimental Plan

Figure 5 shows the simulation setup to validate the RC for short-reach transmission. An OOK signal with 32 GBd and 2^{18} random symbols are transmitted with RRC (0.1 roll-off) as pulse shaping. The digital signal is directly mapped onto the optical domain by assuming an ideal linear transformation. The MZM is assumed ideal to neglect its nonlinear transfer function as a source of degradation and to focus solely on the impact of CD. With that in mind, we only simulated the CD impairment in the SSMF transmission (D = 16.4 ps/nm/km). An additive white Gaussian noise is used to simulate the optical pre-amplifier of the receiver. The noise variance and fiber length (accumulated CD) are swept to meet the target SNR and distance. The AWG is simulated as Gaussian. The signal is then detected by photodetectors that are modeled with an ideal square-law transformation followed by a Bessel filter with the bandwidth matching the electrical bandwidth of the slices. Each of the signals is converted to the digital domain with an ideal transformation and used as an input for the reservoir computing.

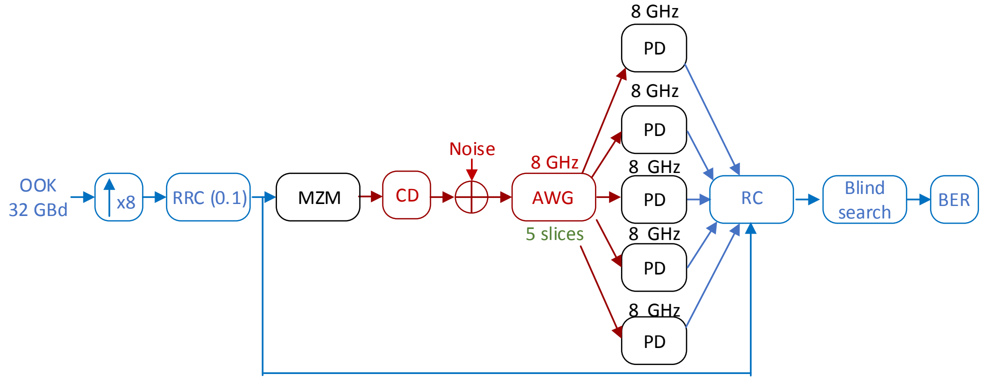

Figure 5: Simulation setup description.

Our preliminary numerical analysis of the setup of Figure 5 already shows significantly transmission reach improvement compared to commonly used techniques such as maximum likelihood sequence estimation or feed-forward neural networks.

The next step of this process is to work on the physical implementation of the concepts studied numerically. In Nokia Bell Labs, we are going to work with an optoelectronic reservoir computing system to validate the simulation results obtained up to now. This experiment is the first step to validate this tool because it is a digital implementation with real numbers. After this first experiment, we will consider extending the work to all-optical or optoelectronic implementation of RC. All-optical and optoelectronic implementation are of particular interest as they can also be applied to NFT systems, where we need to process the complex signal. This is one of the overall goals of the project. We will rely on the experimental testbeds available at Nokia Bell Labs both in terms of short-reach intensity-modulated transmission systems, as well as their coherent NFT-based transmission systems. The following steps are being considered:

1. Apply equalization through digital RC applied to the transmission of on-off keying (OOK) and 4-pulse amplitude modulation (PAM) signals. The experimental waveform digitized after fiber transmission will be equalized through RC and the impact of the reservoir's properties on the signal performance will be studied.
2. The scheme of Figure 2 with an optical pre-processing stage (array waveguide grating or parallel filters) followed by a bank of photodetectors will be implemented. The digitized outputs of the detectors will be injected into the reservoir and the signal quality will be analyzed.
3. More complex optical pre-processing stages will be considered, including nonlinear transformations, e.g., through saturation in semiconductor optical amplifiers or injection-locking of laser dynamics.
4. The extension of the equalization scheme for coherent transmission, including both conventional signaling schemes and nonlinear frequency division multiplexing (NFDM) signals will be considered. This research direction will be pursued in parallel to the characterization of intensity-modulated systems using direct detection. In particular, we will consider RC to provide equalization and phase noise mitigation for NFDM transmission using discrete spectral modulation.

4 ESR 4

There is a high research interest in recent years in applying deep learning techniques to optimize communication systems. In some cases, a specific part of communication system such as coding, pre-equalization or post-equalization is optimized using deep learning and artificial neural networks (ANN). To name some, a low-complexity fiber nonlinearity compensation block was designed using ANNs in [9,10]. In [37], DNN is used to optimize the constellation format for encoding data in long-haul optical transmission when a bitwise soft-decoder is used at receiver. ANNs have also been considered for the equalization module in short reach optical access networks [38]. However, the approach of designing and optimizing the communication system independently on a module-by-module basis can be sub-optimal in terms of the end-to-end performance. The other direction is to optimize the end-to-end performance of a system. It is proposed in [39,40] to optimize the complete communication system from the transmitter to the receiver. Such autoencoder systems, implemented as a single deep neural network, have the potential to achieve the optimal end-to-end performance.

End-to-end optimization based on DNN has gained a recent high interest in communication scenarios in which the optimal communication technique is not known or prohibitive due to complexity reasons. Due to the dispersive and nonlinear nature of optical fiber, optical fiber transmission is an example of such scenarios. Different methods for the end-to-end optimization have been investigated for so-called short reach IM/DD optical systems in [8,41-44]. It has been shown that those optimized systems can reach much higher bit rate and spectral efficiency than the classical systems. However, there are no publication yet, to the best of our knowledge, for end-to-end optimization of high-performance long-haul coherent optical fiber systems.

4.1 Research and Experimental Plan

We are investigation the DNN solutions for end-to-end optimization of long-haul coherent optical fiber systems. In the first step as a proof of concept, we are working on modelling nonlinear Schrödinger equation (NLSE) by a DNN. The NLSE is a basic model for an optical fiber. Next, we will verify our DNN models experimentally using the existing coherent system in Nokia Bell Labs. We will next take the following steps:

1. We improve the current End-to-end solutions in terms of complexity and convergence. It means we will test and compare the new algorithms for DNN optimization. For instance, most of the algorithms were developed for DNN with real valued weights while in communication systems, we are working with complex valued signal. We are investigating the possible further gains if we design complex-valued DNN instead of real-valued DNN.
2. We propose new end-to-end solutions based on different types of DNN and applying signal transformations like nonlinear Fourier transform. After verification in simulation, the plan is to test our proposal in experimental loop of Nokia Bell Labs.

5 Summery

In this deliverable, we described the practical challenges of interest that the ESRs have been planned to tackle in the scope of FONTE. We presented the research direction and the following experimental plans for each ESRs. The experimental plans are defined based on the progress of ESRs in their research.

6 REFERENCES

[1] Govind P Agrawal. *Nonlinear Fiber Optics.* Academic Press, Boston, fifth edition, 2013.

[2] Md Saifuddin Faruk and Seb J Savory. Digital signal processing for coherent transceivers employing multilevel formats. *Journal of Lightwave Technology*, 35(5):1125-1141, 2017.

[3] Andrew D Ellis, Jian Zhao, and David Cotter. Approaching the non-linear Shannon limit. *Journal of Lightwave Technology*, 28(4):423{433, 2009.

[4] Rene-Jean Essiambre, Gerhard Kramer, Peter J Winzer, Gerard J Foschini, and Bernhard Goebel. Capacity limits of optical fiber networks. *Journal of Lightwave Technology*, 28(4):662-701, 2010.

[5] Polina Bayvel, Robert Maher, Tianhua Xu, Gabriele Liga, Nikita A Shevchenko, Domanic Lavery, Alex Alvarado, and Robert I Killey. Maximizing the optical network capacity. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* 374:20140440, 2016.

[6] Ronen Dar and Peter J Winzer. Nonlinear interference mitigation: Methods and potential gain. *Journal of Lightwave Technology*, 35(4):903-930, 2017.

[7] John C Cartledge, Fernando P Guiomar, Frank R Kschischang, Gabriele Liga, and Metodi P Yankov. Digital signal processing for fiber nonlinearities. *Optics express*, 25(3):1916-1936, 2017.

[8] Boris Karanov, Mathieu Chagnon, Felix Thouin, Tobias A Eriksson, Henning Buelow, Domanic Lavery, Polina Bayvel, and Laurent Schmalen. End-to-end deep learning of optical fiber communications. *Journal of Lightwave Technology*, 36(20):4843-4855, 2018.

[9] Christian Hager, Henry D Pfister, Rick M Buetler, Gabriele Liga, and Alex Alvarado. Revisiting multi-step nonlinearity compensation with machine learning. *arXiv preprint arXiv:1904.09807*, 2019.

[10] Christian Hager and Henry D Pfister. Nonlinear interference mitigation via deep neural networks. *In 2018 Optical Fiber Communications Conference and Exposition (OFC),* pages 1-3. IEEE, 2018.

[11] Shaoliang Zhang, Fatih Yaman, Kohei Nakamura, Takanori Inoue, Valey Kamalov, Ljupcho Jovanovski, Vijay Vusirikala, Eduardo Mateo, Yoshihisa Inada, and Ting Wang. Field and lab experimental demonstration of nonlinear impairment compensation using neural networks. *Nature communications*, 10(1):1{8, 2019.

[12] Akira Hasegawa and Takayuki Nyu. Eigenvalue communication. *Journal of lightwave technology*, 11(3):395-399, 1993.

[13] Mansoor I Yousefi and Frank R Kschischang. Information transmission using the nonlinear fourier transform, part i-iii: Spectrum modulation. *IEEE Transactions on Information Theory*, 60(7):4346{4369, 2014.

[14] Jaroslaw E Prilepsky, Stanislav A Derevyanko, Keith J Blow, Ildar Gabitov, and Sergei K Turitsyn. Nonlinear inverse synthesis and eigenvalue division multiplexing in optical fiber channels. *Physical review letters*, 113(1):013901, 2014.

[15] Mansoor Isvand Yousefi. *Information transmission using the nonlinear Fourier transform.* PhD thesis, University of Toronto, 2013.

[16] Sergei K Turitsyn, Jaroslaw E Prilepsky, Son Thai Le, Sander Wahls, Leonid L Frumin, Morteza Kamalian, and Stanislav A Derevyanko. Nonlinear fourier transform for optical data processing and transmission: advances and perspectives. *Optica*, 4(3):307-322, 2017.

[17] Son Thai Le, Vahid Aref, and Henning Buelow. Nonlinear signal multiplexing for communication beyond the Kerr nonlinearity limit. *Nature Photonics*, 11(9):570, 2017.

[18] Son Thai Le. Nonlinear frequency division multiplexing: from single polarization to dual polarization. In *2019 24th OptoElectronics and Communications Conference (OECC) and 2019 International Conference onPhotonics in Switching and Computing (PSC*), pages 1-3. IEEE, 2019.

[19] Shabat, A., Zakharov, V.: 'Exact theory of two-dimensional self- focusing and one-dimensional self-modulation of waves in nonlinear media', Soviet Phys. JETP, 1972, 34, (1), pp. 62–69.

[20] Son Thai Le, Jaroslaw E. Prilepsky, and Sergei K. Turitsyn, "Nonlinear inverse synthesis technique for optical links with lumped amplification," Opt. Express 23, 8317-8328 (2015).

[21] Tao Gui, Gai Zhou, Chao Lu, Alan Pak Tao Lau, and Sander Wahls. Nonlinear frequency division multiplexing with b-modulation: Shifting the energy barrier. *Optics express*, 26(21):27978-27990, 2018.

[22] Tao Gui, Terence H Chan, Chao Lu, Alan Pak Tao Lau, and Ping-Kong Alexander Wai. Alternative decoding methods for optical communications based on nonlinear Fourier transform. *Journal of Lightwave Technology*, 35(9):1542-1550, 2017.

[23] Rasmus T Jones, Simone Gaiarin, Metodi P Yankov, and Darko Zibar. Time-domain neural network receiver for nonlinear frequency division multiplexed systems. *IEEE Photonics Technology Letters*, 30(12):1079-1082, 2018.

[24] Oleksandr Kotlyar, Morteza Kamalian Kopae, Jaroslaw E Prilepsky, Maryna Pankratova, and Sergei K Turitsyn. Machine learning for performance improvement of periodic NFT-based communication system. In *2019 European Conference on Optical Communications*, 2019.

[25] Morteza Kamalian-Kopae, A Vasylchenkova, O Kotlyar, M Pankratova, J Prilepsky, and S Turitsyn. Artificial neural network-based equaliser in the nonlinear Fourier domain for fibre-optic communication applications. In *2019 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC)*, pages 1-1. IEEE, 2019.

[26] Jonas Koch, Rebekka Weixer, and Stephan Pachnicke. Equalization of soliton transmission based on nonlinear Fourier transform using neural networks. In *45th European Conference on Optical Communication (ECOC 2019)*, 2019.

[27] Benedikt Leible, Francisco Javier Garcia-Gomez, and Norbert Hanik. Clustering algorithm for detection in the discrete nonlinear Fourier spectrum. In *45th European Conference on Optical Communication (ECOC 2019),* 2019.

[28] Bajaj et al., " Exact Nonlinear Frequency Division Multiplexing in Lossy Fibers", accepted in Euro. Conf. Opt. Commun., Dublin, Ireland, 2019.

[29] Serkin, V. N., Belyaeva, T. L.: 'Optimal control of optical soliton parameters: Part 1. The Lax representation in the problem of soliton management', Quantum Electronics, 2001, 31, (11), pp 1007-1015.

[30] Stenz, A. J., Boyd, R. W., Evans, A. F.: 'Dramatically improved transmission of ultrashort solitons through 40 km of dispersion-decreasing fiber', Opt. Lett., 1995, 20, (17), pp 1770-1772.

[31] Richardson, D. J., Chamberlin, R. P., Dong, L., et. al.: 'High-quality, soliton loss-compensation in a 38km dispersion decreasing fibre', Electronics Letters, 1995, 31 (19), pp 1681-1682.

[32] G. Khanna, B. Spinnler, S. Calabrò, E. De Man and N. Hanik, "A Robust Adaptive Pre-Distortion Method for Optical Communication Transmitters," in IEEE Photonics Technology Letters, vol. 28, no. 7, pp. 752-755, 1 April1, 2016.

[33] A. Rezania, J. C. Cartledge, A. Bakhshali and W. -. Chan, "Compensation Schemes for Transmitter- and Receiver-Based Pattern-Dependent Distortion," in IEEE Photonics Technology Letters, vol. 28, no. 22, pp. 2641-2644, 15 Nov.15, 2016.

[34] P. W. Berenguer et al., "Nonlinear Digital Pre-distortion of Transmitter Components," in Journal of Lightwave Technology, vol. 34, no. 8, pp. 1739-1745, 15 April15, 2016.

[35] Dan Sadot, Yaron Yoffe, Hananel Faig, and Eyal Wohlgemuth, "Digital Pre-Compensation Techniques Enabling Cost-Effective High-Order Modulation Formats Transmission," J. Lightwave Technol. 37, 441-450 (2019)

[36] Changsoo Eun and E. J. Powers, "A new Volterra predistorter based on the indirect learning architecture," in IEEE Transactions on Signal Processing, vol. 45, no. 1, pp. 223-227, Jan. 1997.

[37] D. Zibar, M. Piels, R. Jones, and C. Schaffer, “Machine learning tech- ¨niques in optical communication,” J. Lightw. Technol., vol. 34, no. 6, pp. 1442–1452, Mar. 2016.

[38] V. Houtsma, E. Chou, and D. van Veen, “92 and 50 Gbps TDM-PON using neural network enabled receiver equalization specialized for PON,” in Proc. OFC 2019.

[39] T. O’Shea and J. Hoydis, "An Introduction to Deep Learning for the Physical Layer," in IEEE Transactions on Cognitive Communications and Networking, vol. 3, no. 4, pp. 563-575, Dec. 2017.

[40] S. Doerner, S. Cammerer, J. Hoydis and S. ten Brink, “Deep learning-based communication over the air,” IEEE J. Sel. Topics Signal Process., vol. 12, no. 1, pp. 132-143, Feb. 2018.

[41] M. Chagnon, B. Karanov and L. Schmalen, “Experimental demonstration of a dispersion tolerant end-to-end deep learning-based IM-DD transmission system,” in Proc. ECOC, Rome, Italy, Sep. 2018.

[42] B. Karanov, D. Lavery, P. Bayvel and L. Schmalen, “End-to-end optimized transmission over dispersive intensity-modulated channels using bidirectional recurrent neural networks,” Opt. Express, vol. 27, no. 14, pp. 19650-19663, 2019

[43] B. Karanov, M. Chagnon, V. Aref, D. Lavery, P. Bayvel, and L. Schmalen,” Concept and Experimental Demonstration of Optical IM/DD End-to-End System Optimization using a Generative Model” in Proc. OFC 2020.

[44] B. Karanov, G. Liga, V. Aref, D. Lavery, P. Bayvel, and L. Schmalen. "Deep Learning for Communication over Dispersive Nonlinear Channels: Performance and Comparison with Classical Digital Signal Processing." In *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 192-199. IEEE, 2019.