Sparse Identification for Nonlinear Optical Communication Systems

Mariia Sorokina*, Stylianos Sygletos, and Sergei Turitsyn

Aston Institute of Photonic Technologies, Aston University, B4 7ET Birmingham UK

*Corresponding author: m.sorokina@aston.ac.uk

ABSTRACT

We have developed a low complexity machine learning based nonlinear impairment equalization scheme and demonstrated its successful performance in SDM transmission links achieving compensation of both inter- and intra- channel Kerr-based nonlinear effects. The method operates in one sample per symbol and in one computational step. It is adaptive, i.e. it does not require a knowledge of system parameters, and it is scalable to different power levels and modulation formats. The method can be straightforwardly expanded to multi-channel systems and to any other type of nonlinear impairment.

Keywords: nonlinear analysis, machine learning, fiber optic communications, spatial division multiplexing.

1. INTRODUCTION

Demand for data is high with today's online culture and the introduction of 8K TV, the Internet of Things and the ever-increasing use of streaming services mean that this demand could outstrip network capacity. Novel techniques, such as machine learning, could help to future-proof our broadband infrastructure.

Here we propose the use of sparse identification method for optical systems, or SINO. This new approach determines the optimum number of variables in the transmission system required for adaptive mitigation of effects (nonlinearities in fibre optic cable) that limit the throughput of standard optical fibre. The SINO method is significantly less complex than other similar compensation techniques. This bodes well for future commercial deployment. SINO is particularly useful for flexible smart-grid networks, as it does not require a knowledge of system parameters and is scalable to difference power levels. Such networks are more sustainable and more reliable, considering the needs of modern society.

Up to date, non-linear impairment mitigation has been considered mostly for single mode fiber links with a number of techniques being proposed both in the optical and electronic domain [1-4]. Most of the research in the field of electronic compensation has been focused on back-propagation algorithms that emulate optical transmission in the digital domain through an inverse fiber link (with reversed order segments and opposite sign parameters), realized either by means of a Split-Step Fourier method [10,11] or Volterra Series Transfer Functions [5-7]. Despite various simplifications that have been proposed, both approaches are considered to be highly complex because they require multiple computational steps along the link and at least two samples per symbol. This has heavily discouraged any effort for future commercial deployment even for legacy SMF fiber systems.

On the other hand, perturbation analysis of the Manakov equations has led to the development of efficient equalization methods that can mitigate accumulated intra-channel impairments in a single computational step and one sample-per-symbol [8-9]. Central to this approach has been the identification of the perturbation coefficients that describe the interaction of each symbol with its preceding and succeeding symbols in the transmission channel [10-13]. For static connections and specific pulse shapes, such as Gaussian or sinc, the perturbation coefficients can be derived analytically and stored in a look-up table [8, 14]. Since the total number of terms can be excessively high for dispersion un-managed links, where the channel memory is long, exploiting common symmetries and quantizing the coefficients are two of the techniques that have been recently proposed for complexity reduction.

Furthermore, to achieve operation in reconfigurable network environments, an adaptive version of the method that uses training sequences and decision-directed least-mean squares algorithms has been introduced in [15]. This enables to identify the perturbation coefficients before establishing any new connection and without prior knowledge of the corresponding transmission link parameters. The latter progress places the perturbation method in a broader context and signifies its practical importance in future optical networks.

With this paper, we expand the application of perturbation-based nonlinear compensation in few mode fiber transmission systems by introducing a novel channel model, which captures nonlinear interplay between different modes in weak and strong coupling regime: Sparse Identification for Nonlinear Optical communication systems: SINO method. Contrary to the aforementioned approaches that deal only with the intra-channel nonlinearities, here, the proposed SINO method takes into account also the nonlinear interaction between the co-propagating modes by introducing additional inter-channel perturbation terms. The associated complexity scaling was addressed by adapting sparse identification method [16], which makes use of the Lasso algorithm [17-18], thus, enabling computation of perturbation coefficients with inherent principal component analysis. The latter minimizes a mean squared error (MSE) estimation based on the training sequence and removing redundant predictors to improve model accuracy. The method does not require knowledge of the transmission line and is applicable for multi-dimensional multichannel systems.

2. METHOD

It is common to use correlations for signal analysis, as it enables to reveal the fundamental relations between the components. Here we are interested in representing each received sample dependent on its neighbours via a clear closed-form expression: $Y_k = \mathbf{\Theta}(X)\mathbf{\Xi}$, here in bold matrices and vectors are denoted. Thus we express the output symbol Y at each time slot k via its neighbours using $\mathbf{\Xi}$ – an interaction matrix. Note that due to nonlinear mixing we consider various nonlinear combinations of samples, thus we create an appropriate alphabet matrix $\mathbf{\Theta}(X) = [\ldots X_k \ X_k^2 X_{k+1}^* \ldots]$. As there is dependence only between small number of components, it means that only if we take a sweep of the neighbours only a few components in $\mathbf{\Theta}(X)$ will contribute, consequently the matrix $\mathbf{\Xi}$ is sparse. The latter enables to make use of sparsity, in particular, use recently proposed sparse identification in nonlinear dynamical systems (SINDY) [16].

Finally, the SINO method is scalable to different modulation formats and signal powers as the matrix $\mathbf{\Xi}$ remains unchanged, while power and modulation format influence are naturally incorporated in $\mathbf{\Theta}$. Thus, once the matrix $\mathbf{\Xi}$ has been identified during the establishment of a connection, and as long as the memory properties of the channel do not change, there is no need for retraining the algorithm and the same $\mathbf{\Xi}$ can be used for any other modulation format (see Fig. 1a). This property makes the method extremely useful for flexible smart-grid network applications.

Overall procedure is as follows (see Fig. 1b):

1) We start by choosing the sample of interest Y_k. Then we identify the a number of neighboring components and list all possible nonlinear combinations of those to take into account χ_3 nonlinearity – $\mathbf{\Theta}(X)$

2) We apply regression algorithm, e.g. Lasso, to identify the interaction matrix $\mathbf{\Xi}$. Note, the matrix needs to be sparse. We choose the values of $\mathbf{\Xi}$ that correspond to sparsity knob which satisfies the minimum mean squared error criteria (see Fig. 1b). Once the matrix is identified the problem is solved;

3) We verify the solution by plotting the measured and predicted values for a different data set alongside each other and the corresponding prediction error.
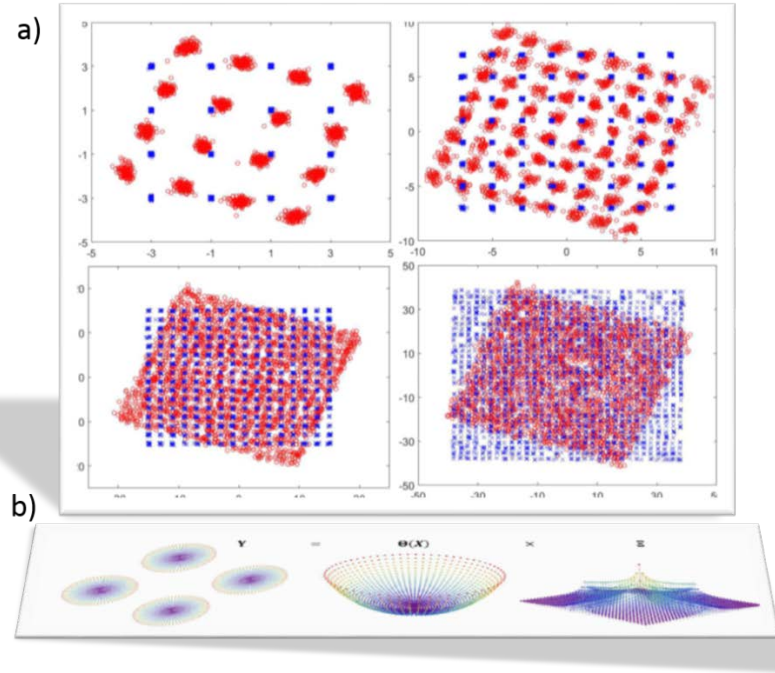

Figure 1: a) Demonstration of the method efficiency for uncompensated (red) and recovered (blue) constellations for 16-, 64-, 256-, and 1024- QAM after 1000km of single mode transmission with average launched power of -3dBm; b) Schematic representation of the principle: the output signal $\mathbf{Y}$ is represented through nonlinear combination of input signal $\mathbf{\Theta}(X)$ via sparse matrix $\mathbf{\Xi}$.

2. RESULTS

The simulations parameters are summarized in Table 1. We investigated the transmission of a single wavelength spatial super-channel, along a 10-span FMF link of total length 10×100 km = 1000 km. We considered D spatial modes, each one of them including two polarization states, with D taking values of 1, 3, and 6. On each of the 2D subchannels we launched a 4096-symbol, 16-QAM modulated stream of root-raised cosine pules (0.01 roll-off) at a symbol rate of 32 GBaud and sampling rate of 16 samples-per-symbol (SpS). A FM-EDFA of 4.5 dB noise figure was considered after each span for compensating the propagation losses. Our focus was on investigating the equalization performance on the Kerr nonlinear effects only. Therefore, for the signal

transmission we considered for every mode a Manakov-type of the propagation equation, where the nonlinear effects are already averaged over all polarization states due to the randomly changing birefringence. Also, for simplicity and without loss of generality, the modes belonged in the same group and they were strongly coupled, which implies a single nonlinear coefficient both for intra- and inter-modal effects (i.e. 1.4 1/W/km), and common group velocity and chromatic dispersion parameters. Finally, the performance was compared in terms of Q^2-factor which as calculated from the error vector magnitude (EVM), according to $Q^2 = 1/EVM^2$. In Fig. 2, we compare the equalization performance of our proposed technique (dotted lines) with the case of equalizing linear impairments only (dashed lines), as well as, with the case of having ideal digital back propagation (DBP) (solid lines). Ideal DBP endures the full compensation of deterministic non-linearity and it is limited by nonlinear signal-noise interactions The proposed nonlinear equalization scheme, results in an above 3 dB improvement for single mode transmission, which decreases to 2 and 1 dB for the 3 and 6 mode cases, respectively. The decrease in performance is attributed to the fact that we had considered only the 1st order terms in the nonlinear expansion of the perturbation model. Our proposed method can be straightforwardly applied to higher orders by augmenting the library matrix Θ with higher order term of X, however, this is an issue to be addressed in a future work. The performance of the proposed algorithm depends strongly on the calculation accuracy through the evaluation of the sparse matrix Ξ. The calculations can be further improved by using more advanced methods for sparse matrix calculation than LASSO.

Table 1: System parameters

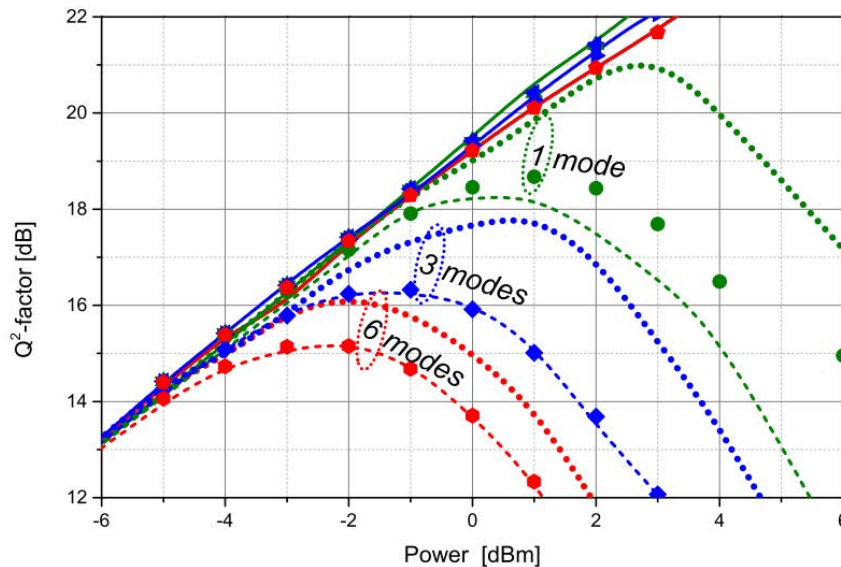|	
Bandwidth	32 GBaud
RRC Roll off	0.01
Distance	10×100 km
Noise Figure	4.5 dB
Attenuation coefficient	0.2 dB/km
CD	17 ps/nm/km
Nonlinearity	1.4 1/W/km

Figure 2. Q^2-factor vs. Launch power per spatial mode (i.e. average power of two pol. multiplexed sub-channels) using the proposed method (dotted) compared to ideal digital back propagation (solid) and linear compensation (dashed), for 1, 3 and 6 spatial modes (green, blue, red). Analytic expressions based on Gaussian noise model are plotted by filled symbols.

Yet, in this simple configuration, which operated with a single sample per symbol and required just a single matrix multiplication, we were able to outperform conventional compensation techniques such as DBP.

An important advantage of the proposed algorithm is that it ensures lowest complexity by evaluating the coupling matrix elements while removing the redundant terms. For example, in the simulated case the number of unique non-zero elements was 48 compared to the estimation of lowest complexity previous algorithm with adaptive filtering [15] 96 components $(M - 1)\ln(M - 1) + 3M - 1$ and non-optimized 324 $(2M + 1)^2$ [8], where $M = \lfloor B^2\beta_2 L/2 \rfloor$ - estimation on channel memory ($\lfloor x \rfloor$ denotes a floor function of a variable x). Furthermore, the previous algorithms were developed only for compensation of inter-channel nonlinearities, while in multichannel operation the complexity will increase proportionally to the number of channels. And the proposed algorithm with the inherent principal component analysis is crucial. In comparison using conventional

DBP with 2 samples per symbol and single step per span calculation requires about 1550 operations, while the performance remains the same.

4. CONCLUSIONS

We have developed a low complexity machine learning based nonlinear impairment equalization scheme and demonstrated its successful performance in SDM transmission links achieving compensation of both inter- and intra- channel Kerr-based nonlinear effects. The method operates in one sample per symbol and in one computational step. It is adaptive, i.e. it does not require a knowledge of system parameters, and it is scalable to different power levels and modulation formats. Finally, although it has been developed for single wavelength spatial super-channels it can be straightforwardly expanded to multi-channel systems and to any other type of nonlinear impairment.

ACKNOWLEDGEMENTS

REFERENCES

[1] M. A. Sorokina and S. K. Turitsyn, "Regeneration limit of classical Shannon capacity," *Nat. Commun.* 5, 3861 (2014).

[2] E. Temprana, E. Myslivets, B.P.-P. Kuo, L. Liu, V. Ataie, N. Alic, and S. Radic, "Overcoming Kerr-induced capacity limit in optical fiber transmission," *Science* 348, 1445-1448 (2015).

[3] E. Ip, "Nonlinear compensation using backpropagation for polarization-multiplexed transmission," *J. Lightwave Technol.* 28(6), 939-951 (2010).

[4] S. J. Savory, "Digital coherent optical receivers: algorithms and subsystems," *IEEE J. Sel. Top. Quantum Electron.* 16(5), 1164-1179 (2010).

[5] K. Peddanarappagari and M. Brandt-Pearce, "Volterra series transfer function of single-mode fibers," *J. Lightw. Technol.*, 15(12), 2232-2241 (1997).

[6] M. Schetzen, "*The Volterra and Wiener Theories of Nonlinear Systems,*" Krieger (2006).

[7] A. Amari, P. Ciblat, andY. Jaouen, "Fifth-order Volterra series based nonlinear equalizer for long-haul high data rate optical fiber communications," in *Proc. Asilomar Conference ACSSC* (2014).

[8] Z. Tao, L. Dou, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussen, "Multiplier-free intrachannel nonlinearity compensating algorithm operating at symbol rate," *J. Lightwave Technol.* 29(17), 2570-2576 (2011).

[9] P. Johannisson and M. Karlsson, "Perturbation analysis of nonlinear propagation in a strongly dispersive optical communication system," *J. Lightwave Technol.* 31(8), 1273-1282 (2013).

[10] Y. Gao, J. C. Cartledge, A. S. Karar, S. S. Yam, M. O'Sullivan, C. Laperle, A. Borowiec, and K. Roberts, "Reducing the complexity of perturbation based nonlinearity pre-compensation using symmetric EDC and pulse shaping," Opt. *Express* 22(2), 1209-1219 (2014).

[11] Q. Zhuge, M. Reimer, A. Borowiec, M. O'Sullivan, and D. V. Plant, "Aggressive quantization on perturbation coefficients for nonlinear pre-distortion," in *Proc. Optical Fiber Communications Conference and Exhibition* (Optical Society of America), paper Th4D.7 (2014).

[12] Z. Li, W.-R. Peng, F. Zhu, and Y. Bai, "Optimum quantization of perturbation coefficients for perturbative fiber nonlinearity mitigation," in *Proc. European Conference on Optical Communication*, paper We.1.3.4 (2014).

[13] Z. Li, W. Peng, F. Zhu, and Y. Bai, "MMSE-based optimization of perturbation coefficients quantization for fiber nonlinearity mitigation," *J. Lightwave Technol.* 33(20), 4311-4317 (2015).

[14] A. Ghazisaeidi and R.-J. Essiambre, "Calculation of coefficients of perturbative nonlinear pre-compensation for Nyquist pulses," in *Proc. European Conference on Optical Communication*, paper We.1.3.3 (2014).

[15] M. Malekiha, I. Tselniker, and D. V. Plant, "Efficient nonlinear equalizer for intra-channel nonlinearity compensation for next generation agile and dynamically reconfigurable optical networks," *Opt. Express* 24, 4097-4108 (2016) .

[16] S. Brunton, J. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proc. Natl. Acad. Sci.* 113, 3932-3937 (2016).

[17] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. R. Stat. Soc. B* 58(1), 267-288 (1996).

[18] T. Hastie, R. Tibshirani, and J. Friedman, "*The Elements of Statistical Learning,*" vol. 2, Springer (2009).