Reduction of Nonlinear Inter-Subcarrier Intermixing in Coherent Optical OFDM by a Fast Newton-based Support Vector Machine Nonlinear Equalizer

Elias Giacoumidis, *Member IEEE, Member OSA*, Sofien Mhatli, Marc F. C. Stephens, Athanasios Tsokanos, Jinlong Wei, *Member IEEE, Member OSA*, Mary E. McCarthy, Nick J. Doran, *Senior Member IEEE, Senior Member OSA,* and Andrew D. Ellis, *Senior Member IEEE, Senior Member OSA*

***Abstract*—A fast Newton-based support vector machine (N-SVM) nonlinear equalizer (NLE) is experimentally demonstrated, for the first time, in 40 Gb/s 16-quadrature amplitude modulated coherent optical orthogonal frequency division multiplexing at 2000 km of transmission. It is shown that N-SVM-NLE extends the optimum launched optical power by 2 dB compared to the benchmark Volterra-based NLE. The performance improvement by N-SVM is due to its ability of tackling both deterministic fiber-induced nonlinear effects and the interaction between nonlinearities and stochastic noises (e.g. polarization-mode dispersion). N-SVM is more tolerant to inter-subcarrier nonlinear crosstalk effects than Volterra-based NLE, especially when applied across all subcarriers simultaneously. In contrast to the conventional SVM, the proposed algorithm is of reduced classifier complexity offering lower computational load and execution time. For a low *C*-parameter of 4 (a penalty parameter related to complexity), an execution time of 1.6 sec is required for N-SVM to effectively mitigate nonlinearities. Compared to conventional SVM, the computational load of N-SVM is ~6 times lower.**

***Index Terms*—Coherent detection, nonlinearity mitigation, support vector machines, coherent optical OFDM.**

I. INTRODUCTION

THE data rate in an optical transmission system is currently limited by amplified spontaneous emission, which determines the minimum power launched into each fiber span, and the interplay between chromatic dispersion (CD) and Kerr fiber nonlinearity, which limits the maximum launch power [1].

Manuscript received December 30, 2016. revised XXXX, 2017; XXXX. This work was supported by the Centre of Excellence (CUDOS, CE110001018); EPSRC (EP/J017582/1, EP/L000091/1, EP/M005283/1); and by Sterlite Technol. Ltd. We thank Dr Son Le from Nokia-Bell-Labs (Germany) and Prof Benjamin Eggleton from the University of Sydney for their guidance.

Elias Giacoumidis is with the University of Sydney, School of Physics, NSW 2006, Sydney, Australia (email: e.giacoumidis@physics.usyd.au).

Sofien Mhatli is with SERCOM-Lab, EPT Université de Carthage, Tunis 2078, Tunisia (email: sofien_mhatli@yahoo.fr).

Marc F. C. Stephens, Mary E. McCarthy, Nick J. Doran, and Andrew D. Ellis are with Aston University, AIPT, B4 7ET, Birmingham, UK (e-mail: {m.stephens, m.mccarthy, n.j.doran, andrew.ellis}@aston.ac.uk).

Athanasios Tsokanos is with the University of Hertfordshire, Hatfield, Hertfordshire, UK (email: a.tsokanos@herts.ac.uk).

Jinlong Wei is with the Huawei Düsseldorf GmbH, European Research Center, 80992 München, Germany (email: jinlong.wei@huawei.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier XX.XX/JLT.2017.XXXXX

To increase the data rate of current-generation coherent systems, fiber nonlinearity compensation is required to enable higher launch powers, thereby providing enough optical signal-to-noise ratio to support larger constellation sizes [2]. State-of-the-art fiber nonlinearity compensators (NLC) include digital signal processing (DSP)-based techniques such as digital back-propagation (DBP) [2], [3], reduced complexity Volterra-based nonlinear equalization (NLE) [4], and phase-conjugated twin-waves [5], which tackle nonlinearities of deterministic nature. However, in coherent long-haul optical systems the interaction between nonlinear phenomena with random noises such as polarization-mode dispersion (PMD) results in stochastic nonlinear distortion, which can be partially mitigated using machine learning in the digital domain such as support vectors machines (SVM) [6]–[10].

On the other hand, coherent optical orthogonal frequency division multiplexing (CO-OFDM) is an excellent candidate for long-haul communications because of its high spectral efficiency, flexibility, and tolerance to chromatic dispersion (CD) and PMD. However, due to its high peak-to-average power ratio the deterministic nonlinear cross-talk effects among subcarriers such as inter-subcarrier intermixing (ICI) cross-phase modulation (XPM) and four-wave mixing (FWM) are significantly enhanced causing an additional "stochastic-like" interference [6], [7]. SVM-based NLEs [6]–[10] have shown promising results in CO-OFDM. Nevertheless, since optimization usually requires many steps to converge (in the order of 30) [7], implementation in real-time processing is impractical.

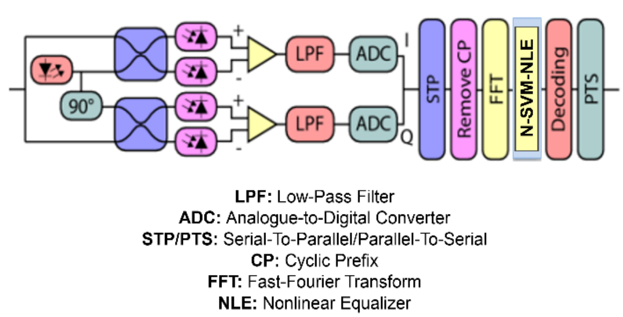

Fig. 1. Block diagram of the CO-OFDM receiver equipped with the proposed N-SVM-NLE.

In this paper, we experimentally demonstrate, for the first time, a fast classification SVM-NLE of reduced classifier complexity using the Newton-method (N-SVM) [11] in 16 quadrature amplitude modulated (16-QAM) CO-OFDM at 40 Gb/s, transmitted at 2000 km of standard single-mode fiber (SSMF). It is shown that compared to the benchmark deterministic Volterra-based NLE, N-SVM extends the optimum launched optical power (LOP) by 2 dB with very low DSP computational load and execution time. N-SVM tackles ICI nonlinear crosstalk effects more effectively than Volterra-NLE especially when applied across all subcarriers simultaneously, rather than on each subcarrier separately.

The paper is organized as follows: Section II analyses the principle of the proposed N-SVM-NLE and the benchmark Volterra-NLE for 16-QAM CO-OFDM. Section III describes the experimental CO-OFDM setup. Section IV presents the experimental results of N-SVM-NLE and Volterra-NLE for CO-OFDM at 2000 km of transmission, and finally in Section V the paper is concluded.

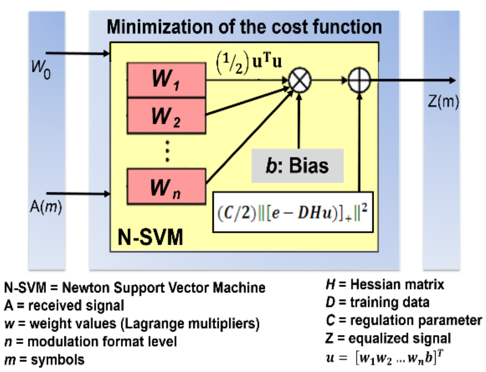

Fig. 2. Block diagram of proposed N-SVM for the adopted single-channel/polarization 16-QAM CO-OFDM receiver.

II. Principle of Newton Support Vector Machine-NLE

A. Operation of N-SVM-NLE for 16-QAM CO-OFDM

In Fig. 1 the block diagram of the CO-OFDM receiver equipped with the N-SVM-NLE is depicted, where the received optical signal is converted back to an electrical one through a homodyne 90^0 coherent detector. Afterwards, OFDM demodulation process follows similarly to [6], where serial-to-parallel (STP), removal of cyclic prefix (CP) and fast Fourier transform (FFT) are processed. After the FFT block the proposed N-SVM-NLE takes place for all subcarriers simultaneously before decoding and parallel-to-serial (PTS) conversion. The proposed N-SVM-NLE implements a fast Newton method that suppresses input space features for a nonlinear programming formulation of supervised SVM classifiers. This stand-alone method can handle classification problems in very high dimensional spaces. An implicit Lagrangian formulation of an SVM classifier which leads to a highly effective iterative scheme [11] is solved in this algorithm by a Newton method which handles classification problems in just a few steps. In order to handle a 16-QAM constellation mapper which has a very large dimensional input, a fast-finite Newton method is employed to find the unconstrained unique global minimum solution of the implicit Lagrangian associated with the classification problem. The solution is obtained by solving a system of nonlinear equations, a finite number of times. The algorithm implements the Newton method with an Armijo step-size [12] and establishes its finite global termination to the unique solution. All vectors are column vectors unless transposed to a row vector by a T superscript. The 2-norm of a vector x is denoted by $\|x\|$. The matrix A[m×n] is related to the A received signal with m complex OFDM symbols in the n-dimensional real space R^m which defines the order of modulation format level (i.e. 16 for 16-QAM) as depicted in Fig. 2.

In Fig. 2 where e is the column vector of value 1, while w, b, are the normal vector (i.e. weights with w_0 being the initialized weight) and the scalar of the hyperplane (bias), respectively. To control the trade-off between minimizing training errors and model complexity we introduce a slack variable z for each training symbol and a "penalty parameter" C (which controls the trade-off between the slack variable penalty and the margin). Similar to [13] the margin maximization formula in the SVM is replaced by the least square 2-norm error, which brings out an unconstrained optimization being solved by the finite "stepless" Newton method. The N-SVM formulation thus requires only solutions of nonlinear equations instead of quadratic programming and simultaneously maximizes the margin and minimizes the error as shown in (1):

$$\min f(w,b,z) = \left(\frac{1}{2}\right)\|w\|^2 + Ce^T z \qquad (1)$$
$$\textit{subject to } \; D(Aw - eb) + z \geq e,$$

where $z \in R^m$ is the non-negative slack vector and $C \in R^1$ is a positive constant (C penalty parameter), both used to tune errors and margin size, while A is the received signal. To perform nonlinear N-SVM, the classification sigmoid function is employed. To change from a linear to a non-linear classifier however, we substitute a kernel evaluation in (1) instead of the original 'dot product'. Recent developments for massive nonlinear SVM algorithms [11] reformulate the classification as an unconstrained optimization. By changing the margin maximization to the minimization of $\left(\frac{1}{2}\right)\|w,b\|^2$ and adding with a least squares 2-norm error, the SVM reformulation with nonlinear kernel leads to:

$$\min f(w,b,z) = \left(\frac{1}{2}\right)\|w,b\|^2 + \left(\frac{C}{2}\right)\|z\|^2 \qquad (2)$$
$$\textit{subject to } \; D(Aw - eb) + z \geq e$$

The formulation of (2) can be rewritten by substituting $z = [e - D(Aw - eb)]_+$ leading to (3):

$$\min f(w,b) = \left(\frac{1}{2}\right)\|w,b\|^2 + \left(\frac{C}{2}\right)\|[e - D(Aw - eb)]_+\|^2 \qquad (3)$$

where $(x)_+$ replaces negative components of a vector x by zeros into the objective function f. By setting $[w_1 w_2 \dots w_n b]^T$ to u and $[A - e]$ to H (which is the Hessian matrix [11]), then the SVM formulation of (3) is rewritten by (4):

$$\min f(w,b) = \left(\frac{1}{2}\right)u^T u + \left(\frac{C}{2}\right)\|[e - DHu)]_+\|^2 \qquad (4)$$

B. The "stepless" N-SVMA algorithm

The adopted N-SVM process is described in Fig. 3 showing the finite "stepless" Newton method which solves the strongly convex unconstrained minimization problem in (4). In most of tested cases [11]–[14] this algorithm has given an optimum solution with a few number of iterations varying from 5 to 8.

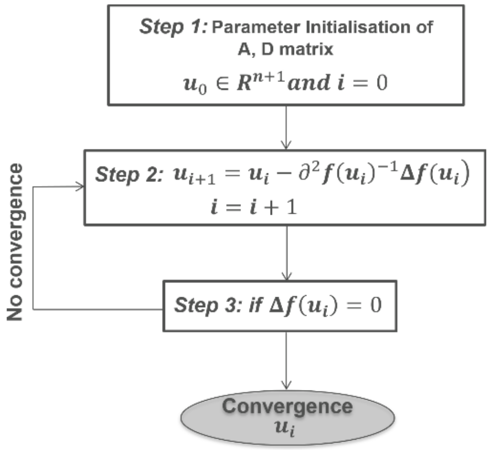

Fig. 3. N-SVM algorithm. A=Received signal; D=training data.

III. BENCHMARK VOLTERRA-NLE FOR 16-QAM CO-OFDM

The adopted Volterra-NLE is similar to [4], accounting for single-band and single-polarization as depicted in Fig. 4. It employs the inverse Volterra-series transfer function (IVSTF) with up to 3rd order Volterra kernels. It should be noted that when higher-order kernels were employed, similar results were revealed [15]. IVSTF-NLE offers ~25% reduced complexity compared to full-step/span DBP [4], [9] and inherits some of the features of the hybrid time-and-frequency domain implementation, such as non-frequency aliasing and simple implementation.

The process of nonlinearity compensation by Volterra-NLE is described as follows: The input OFDM signal is first converted to frequency domain by FFT. The Volterra-NLE compensates CD using a linear compensator.

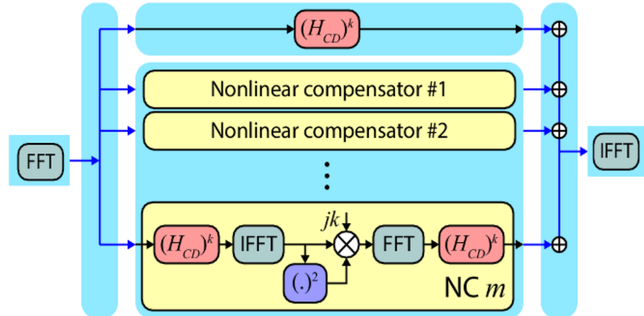

Fig. 4. IVSTF-NLE [4] for 16-QAM CO-OFDM. (I)FFT: (inverse) fast-Fourier transform; H_{CD}: system chromatic dispersion; NC: nonlinearity compensation; *k*: constant related to the nonlinear distortion and the total power. *m*: number of nonlinearity compensators.

On the other hand, the number of required nonlinear compensators depends on the number of homogeneous spans in the transmission link. The output of the linear and nonlinear compensator is combined and converted back to time-domain using the inverse FFT (IFFT). The Volterra-NLE procedure can be described from (5) – (9). Since a reduced complexity 3rd order IVSTF is considered, the kernels $H_1(\omega, z)$ and $H_3(\omega_1, \omega_2, \omega - \omega_1 + \omega_2, z)$ are given by,

$$H_3(\omega z) = e^{-\alpha z/2} e^{-j\omega^2 \beta_2 z/2} \tag{5}$$

$$H_3(\omega_1, \omega_2, \omega - \omega_1 + \omega_2, z) = -\frac{j\gamma}{4\pi^2} H_1(\omega, z) \times \frac{1 - e^{-(\alpha + j\beta_2(\omega_1 - \omega)(\omega_1 - \omega_2))z}}{\alpha + j\beta_2(\omega_1 - \omega)(\omega_1 - \omega_2)}. \tag{6}$$

where ω is the optical frequency and ω_1, ω_2 are the dummy variables acting as parameters and influence the interactions of the lightwaves at different frequency, especially the ICI interaction effects. α is the fiber loss, β_2 is the 2nd order CD parameter and γ accounts for the effect of fiber nonlinearity averaging. For an optically amplified N_{span} fiber link with L_{span} being the span length, the corresponding p^{th} inverse is given by the nonlinear kernels as,

$$K_1(\omega) = H_1^{-1}(\omega) = e^{-j\omega^2 \beta_2 N_{span} L_{span}/2} \tag{7}$$

$$K_3(\omega_1, \omega_2, \omega - \omega_1 + \omega_2) = -\frac{j\gamma}{4\pi^2} K_1(\omega)$$

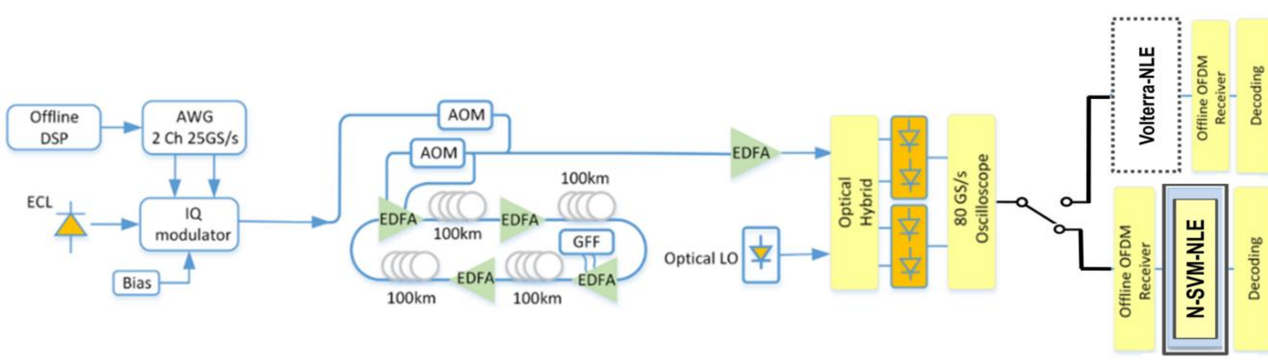

Fig. 5. Experimental setup of 40 Gb/s CO-OFDM equipped with either Volterra-NLE or N-SVM-NLE. ECL: external cavity laser, DSP: digital signal processing, AWG: arbitrary waveform generator, AOM: acousto-optic modulator, EDFA: Erbium-doped fiber amplifier, GFF: gain flatten filter, LO: local oscillator.

$$\times \frac{1-e^{-(\alpha+j\beta_2\Delta\omega)L_{span}}}{\alpha+j\beta_2(\omega_1-\omega)(\omega_1-\omega_2)} \sum_{k=1}^{N_{span}} e^{-jk\beta_2 L_{span}\Delta\omega} \tag{8}$$

$$\approx -\frac{j\gamma}{4\pi^2}\times\frac{1-e^{-\alpha L_{span}}}{\alpha} \times K_1(\omega)\sum_{k=1}^{N_{span}} e^{-jk\beta_2 L_{span}\Delta\omega}. \tag{9}$$

The corresponding compensation scheme representing Eqs. (7) and (9) is applied in Fig. 4. Each nonlinear compensation stage is a realization of

$$K_3(\omega_1,\omega_2,\omega-\omega_1+\omega_2) \approx -\frac{j\gamma}{4\pi^2}\times\frac{1-e^{-\alpha L_{span}}}{\alpha} \times K_1(\omega)e^{-jk\beta_2 L_{span}\Delta\omega}. \tag{10}$$

Finally, since single-polarization is considered we have

$$S_{K_1}(\omega)\iint_{\infty}^{\infty} K_{3,K_1}(\omega_1,\omega_2,\omega-\omega_1+\omega_2) \times A(\omega_1)A^*(\omega_2)\times A(\omega-\omega_1+\omega_2)d\omega_1 d\omega_2 \tag{11}$$

where $S_{K_1}(\omega)$ is derived by passing the received signal through $(H_{CD})^{(K_1,K_2)}$, and nonlinearity compensation is performed by $jk(|.|^2)(.)$ where we multiply the received signal by a constant k related to the nonlinear distortion and the total power. This parameter varies for this configuration and is obtained by sweeping it to get optimum performance, which is part of the calibration of the Volterra-NLE. Finally, the residual CD is compensated passing through $(H_{CD})^{N-(K_1,K_2)}$.

IV. Experimental Setup

Fig. 5 depicts the experimental setup where an external cavity laser (ECL) of 100 kHz linewidth was modulated using a dual-parallel Mach-Zehnder modulator (DP-MZM) in IQ configuration. The DP-MZM was fed with OFDM I-Q components, which was generated offline. The transmission path at 1550.2 nm was a recirculating loop consisting of 20×100 km spans of Sterlite OH-LITE (E) SSMF (attenuation of 18.9-19.5 dB/100 km) controlled by acousto-optic modulator (AOM). The loop switch was located in the mid-stage of the 1st Erbium-doped fiber amplifier (EDFA) and a gain-flattening filter (GFF) was placed in the mid-stage of the 3rd EDFA. The optimum LOP was swept by controlling the output power of the EDFAs. At the receiver, the incoming signal was combined with another 100 kHz linewidth ECL acting as local oscillator. After down-conversion, the baseband signal was sampled using a real-time oscilloscope operating at 80 GS/s and processed offline in *Matlab*. 400 OFDM symbols were generated using a 512-point IFFT in which 210 subcarriers were modulated using 16-QAM. To eliminate inter-symbol-interference from linear effects, a CP of 2 % was included. For fair comparison among linear equalization (LE), Volterra-NLE and the proposed N-SVM-NLE, the net and raw bit-rate were fixed at ~40 Gb/s and ~46 Gb/s, respectively. The N-SVM training overhead was set at 10 % (optimum value for LE) resulting in a training length of 40 symbols. The offline OFDM demodulator included timing synchronization, frequency offset compensation, channel estimation and equalization with the assistance of an initial training sequence, as well as I-Q imbalance and CD compensation using an overlapped frequency domain equalizer employing the overlap-and-save method. When N-SVM-NLE was performed, the LE was neglected due to N-SVM ability of compensating both linear and nonlinear inter-subcarrier crosstalk effects. The CO-OFDM transceiver and transmission parameters are depicted on Table I. The NLEs performances were assessed by Q-factor measurements averaging over 10 recorded traces (~10^6 bits), which was estimated from the bit-error-rate (BER) obtained by error counting after hard-decision decoding. The Q-factor is related to BER by Q = $20\log_{10}[\sqrt{2}erfc^{-1}(2BER)]$. For 16-QAM, a BER of 10^{-3} (forward-error-correction-limit, FEC-limit) results in a Q-factor of ~9.8 dB.

Table I. CO-OFDM transceiver and transmission parameters

Parameter	Value
Net bit-rate (LE, NLEs)	~40 Gb/s
Raw bit-rate (LE, NLEs)	~46 Gb/s
Signal modulation format	16-QAM
OFDM symbols	400
Modulated OFDM subcarriers	210
Cyclic prefix (CP) length	2 %
FFT/IFFT size	512
N-SVM Training overhead	10 %
N-SVM Training symbol length	40 symbols
ECL linewidth	100 KHz
OH-LITE (E) SSMF attenuation	18.9–19.5 dB/100 km
Span number	20
Span length	100 km
Transmission wavelength	1550.2 nm

V. Results and Discussion

In Fig. 6 the Q-factor against the training overhead of N-SVM-NLE is depicted for 16-QAM CO-OFDM at 2000 km of transmission for a LOP of 2 dBm, which is the optimum LOP of LE. It should be noted that changing the training overhead, the raw bit-rate was adjusted accordingly. From Fig. 6 it is evident that a minimum 10 % of training data is required for N-SVM-NLE to effectively tackle the OFDM inter-subcarrier crosstalk effects (e.g. ICI-XPM/FWM). In this paper, 10 % of training data are employed for N-SVM-NLE in all sections.

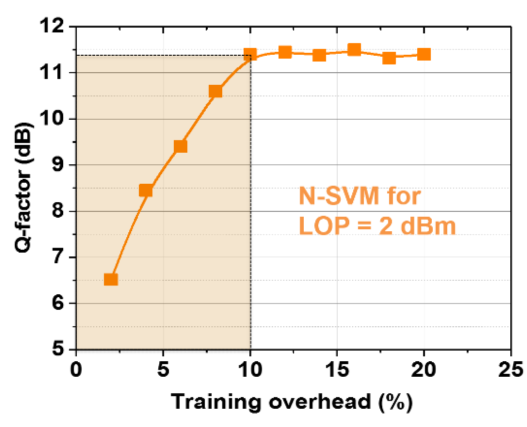

Fig. 6. Q-factor vs. training overhead of N-SVM-NLE for 16-QAM CO-OFDM at 2000 km of transmission for a launched optical power (LOP) of 2 dBm, which is the optimum LOP of linear equalization (LE).

In Fig. 7, the Q-factor against the LOP is plotted for the 40 Gb/s CO-OFDM system at 2000 km of transmission for LE, Volterra-NLE, and N-SVM-NLE. It is shown that compared to Volterra-NLE, the proposed N-SVM-NLE can extend the optimum LOP by 2 dB (FEC-limit at ~9.8 dB), while in comparison to LE it can extend the LOP by ~3.5 dB. To corroborate the N-SVM-NLE performance enhancement, Fig. 8 is plotted, showing the received 16-QAM constellations diagrams for the three types of equalization and without equalization at 6 dBm of LOP.

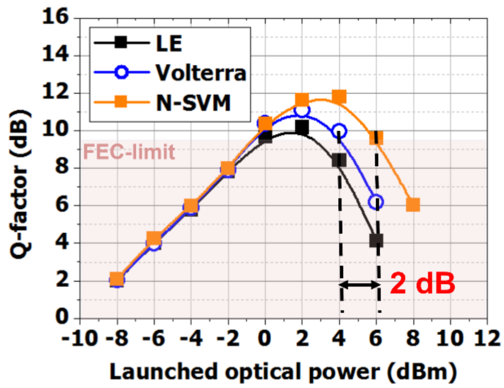

Fig. 7. Q-factor vs. LOP for 16-QAM CO-OFDM when performing LE, Volterra-NLE, and N-SVM-NLE.

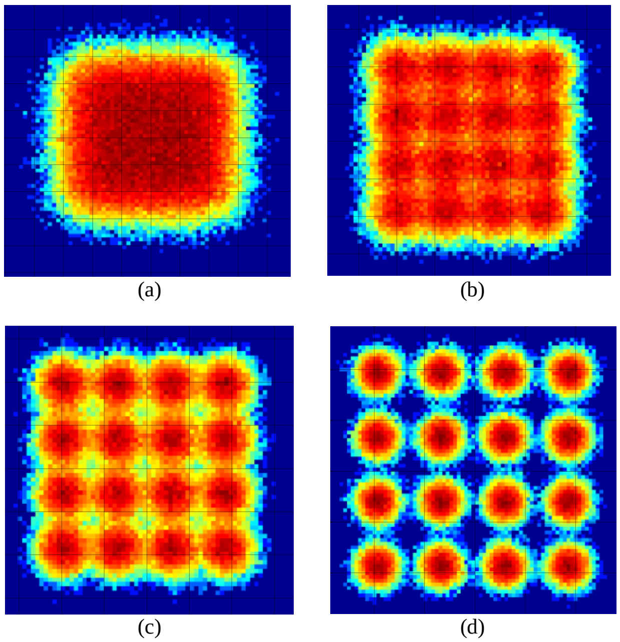

Fig. 8. Received 16-QAM constellation diagrams of CO-OFDM at 2000 km of transmission when the LOP is 6 dBm for the following cases: (a) without equalization, (b) LE, (c) Volterra-NLE, and (d) N-SVM-NLE.

In Fig. 9, the Q-factor against the *C*-parameter (the *C* value from (2)) is plotted for the CO-OFDM system under test at a LOP of 4 dBm. The *C*-parameter (also called "penalty parameter") is related to the computational complexity of N-SVM. It is shown that a *C* of only 4 is required at an execution time of 1.6 sec for stable optimum performance. This time required by the training process is considered for a general-purpose CPU operating at 1.2 GHz. However, this time will be drastically reduced in implementations based on Field-Programmable Gate-Array or Application Specific Integrated Circuits. The minimum required *C* value for N-SVM-NLE is ~6 times less than the corresponding "penalty parameter" of the conventional SVM-NLE reported in [7] for 16-QAM CO-OFDM. This occurs because i) N-SVM performs fast classification tasks that separate cases of different class labels, and ii) the conventional SVM performs both classification and regression analysis in contrast to N-SVM which only classifies the data. It should be noted that a transmission performance comparison between the proposed N-SVM and the conventional SVM [7] is out of the scope of this paper since fair comparison is not feasible.

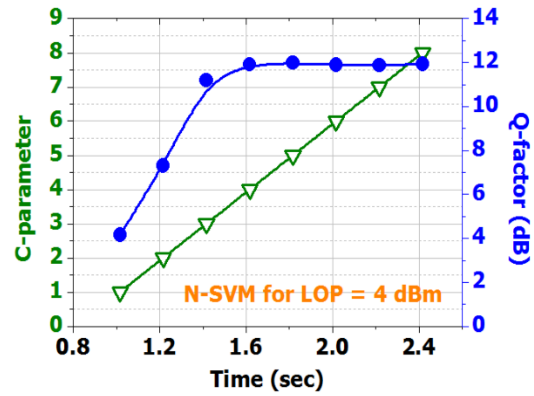

Fig. 9. C-parameter/Time vs. Q-factor for 16-QAM CO-OFDM equipped with N-SVM-NLE at a LOP of 4 dBm.

In Fig. 10 the impact of N-SVM on the nonlinear ICI crosstalk effects is investigated for the adopted CO-OFDM system. A comparison is also made with the benchmark Volterra-NLE to evaluate the impact of stochastic nonlinearities. In Fig. 10, an additional case for exploring the nonlinear phenomena in OFDM is proposed, in which the NLEs under test are performed for each subcarrier. Although this case is unrealistic since it substitutes a separate NLE for each subcarrier, it will provide a holistic and deeper understanding on the physics underlying nonlinear phenomena in CO-OFDM. In Fig. 11, a conceptual diagram is depicted for the application of NLE, and NLE per subcarrier (related to Volterra and N-SVM) on received OFDM signal. N-SVM and Volterra NLEs 'per subcarrier' cases (the dotted lines in Fig. 10) includes 210 NLEs in contrast to the realistic case where 1 NLE process all subcarriers together. In Fig. 10, it is shown that in comparison to the 'per subcarrier' case, when N-SVM is applied across all subcarriers it reduces the fiber nonlinearity penalty by 0.5 dB. This occurs because when applying N-SVM on each subcarrier separately, ICI nonlinear crosstalk effects are not combated. Finally, it is confirmed that CO-OFDM is influenced by stochastic nonlinearities which cannot be tackled by the deterministic Volterra-NLE.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2017.2678511, Journal of Lightwave Technology

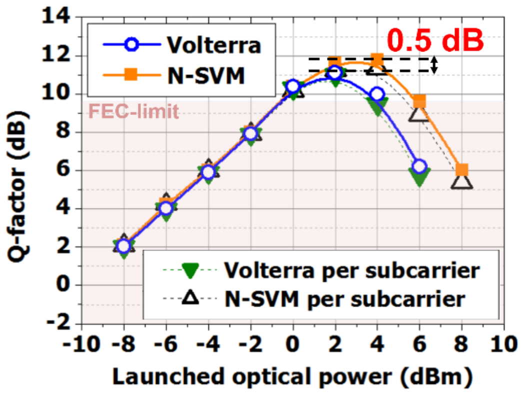

Fig. 10. Q-factor vs. LOP for 16-QAM CO-OFDM equipped with Volterra/N-SVM being processed across all subcarriers and per subcarrier (dotted lines).

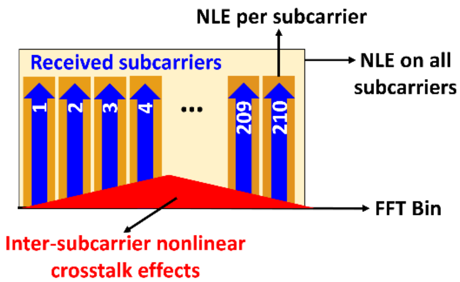

Fig. 11. Conceptual diagram of application of NLE, and NLE per subcarrier (for Volterra and N-SVM) on received OFDM signal.

The results from Fig. 10 indicate that the adopted realistic N-SVM-NLE which accounts for all subcarriers together, provides effective and fast compensation of inter-subcarrier nonlinear crosstalk effects in CO-OFDM.

VI. Conclusion

A novel fast N-SVM-NLE of reduced classifier complexity was experimentally demonstrated in 40 Gb/s 16-QAM CO-OFDM at 2000 km of SSMF. In comparison to Volterra-NLE, the proposed N-SVM extended the optimum LOP by 2 dB with very low computational load and execution time. N-SVM tackled inter-subcarrier nonlinear crosstalk effects more effectively than Volterra-NLE especially when applied across all subcarriers simultaneously.

References

[1] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fibre communications," *Nature*, vol. 411, pp. 1027–1030, April 2001.

[2] D. Rafique, J. Zhao, and A. D. Ellis, "Digital back-propagation for spectrally efficient WDM 112 Gbit/s PM m-ary QAM transmission," *Opt. Exp.*, vol. 19, no.6, pp. 5219–5224, Mar. 2011.

[3] G. Gao, X. Chen, and W. Shieh, "Limitation of fiber nonlinearity compensation using digital back propagation in the presence of PMD," presented at Optical Fiber Communication Conf., Los Angeles, CA, USA, Mar. 4–8, 2016, Paper OM3A.5.

[4] E. Giacoumidis, I. Aldaya, M. A. Jarajreh, A. Tsokanos, S. T. Le, F. Farjady, A. D. Ellis, and N. J. Doran, "Volterra-Based Reconfigurable Nonlinear Equalizer for Coherent OFDM," *IEEE Phot. Technol. Lett.*, vol. 26, no. 14, pp. 1383–1386, July 2014.

[5] X. Liu, A. R. Chraplyvy, P. J. Winzer, R. W. Tkach, and S. Chandrasekhar, "Phase-conjugated twin waves for communication beyond the Kerr nonlinearity limit," *Nature Phot.*, vol. 7, no. 7, pp. 560–568, May 2013.

[6] E. Giacoumidis, S. Mhatli, T. Nguyen, S. T. Le, I. Aldaya, M. E. McCarthy, and B. J. Eggleton, "Kerr-Induced Nonlinearity Reduction in Coherent Optical OFDM by Low Complexity Support Vector Machine Regression-based Equalization," presented at Optical Fiber Communication Conf. (OFC), San Francisco, CA, USA, Mar. 20–24, 2016, Paper Th2A.49.

[7] E. Giacoumidis, S. T. Le, M. E. MacCarthy, A. D. Ellis, and B. J. Eggleton, "Record Intrachannel Nonlinearity Reduction in 40-Gb/s 16QAM Coherent Optical OFDM using Support Vector Machine based Equalization," Australian and New-Zealand Conf. on Optics and Photonics (ANZCOP) and Australian Conf. in Optical Fiber Technologies (ACOFT), Adelaide, Australia, Nov. 29 – Dec. 4, 2015, ISBN: 9781922107664, Postdeadline Paper.

[8] T. Nguyen, S. Mhatli, E. Giacoumidis, L. Van Compernolle, M. Wuilpart, and P. Megret, "Fiber nonlinearity equalizer based on support vector classification for coherent optical OFDM," *IEEE Phot. J.*, vol. 8, no. 2, Feb. 2016.

[9] E. Giacoumidis, S. Mhatli, T. Nguyen, S. T. Le, I. Aldaya, M. E. McCarthy, A. D. Ellis, and B. J. Eggleton, "Comparison of DSP-based nonlinear equalizers for intra-channel nonlinearity compensation in coherent optical OFDM," *Opt. Lett.*, vol. 41, no. 11, pp. 2509–2512, May 2016.

[10] E. Giacoumidis, S. Mhatli, I. Aldaya, M. E. McCarthy, A. D. Ellis, and B. J. Eggleton, "Nonlinear Blind Equalization for 16-QAM Coherent Optical OFDM using Support Vector Machines," presented at European Conf. and Exhibition on Optical Communication (ECOC), Düsseldorf, Germany, Sep. 18–22, 2016, Paper Th.2.P2.

[11] N. Cristianini and J. S.-Taylor, "An introduction to support vector machines and other kernel-based learning methods," Cambridge University Press, Cambridge, UK, Mar. 2000.

[12] L. Armijo, "Minimization of functions having Lipschitz continuous first partial derivatives," *Pacific J. of Mathem.*, vol. 16, no. 1, pp. 1–3, Jan. 1966.

[13] O. L. Mangasarian and D. R. Musicant, "Lagrangian Support Vector Machines," *J. of Machine Learn. Research*, vol. 1, no. 3, pp. 161–177, Jan. 2001.

[14] P. S. Bradley and O. L. Mangasarian, "Feature selection via concave minimization and support vector machines," Machine Learning Proceedings of the 15th International Conf. (ICML), San Francisco, CA, USA, pp. 82–90, July 24–27, 1998.

[15] A. Amari, P. Ciblat, and Y. Jaouën, "Fifth-order Volterra series based nonlinear equalizer for long-haul high data rate optical fiber communications," presented at 48th Asilomar Conf. on Signals, Systems and Computers, Pacific Grove, CA, USA, 2–5 Nov. 2014.

Elias Giacoumidis is a research fellow at the University of Sydney since April 2015. He works for the Terabit per second Photonics flagship project of CUDOS in which he serves as a deputy project leader. He has previously worked for the prestigious Optical Communication Research Groups in Bangor University, Athens Information Technology (AIT), Aston University, and Telecom-ParisTech, where he was also teaching optical system modelling. He has authored and co-authored 90 papers that appeared in international peer-reviewed journals and top conferences being cited 773 times by Google-Scholar with H-factor of 15. His research involves balanced theoretical and experimental exploration in > 40G transmission systems with specialization in key modern DSP modulation techniques (OFDM, CAP, PAM-4 etc.) for local networks, passive optical networks (PONs) and next-generation flexible core networks. Dr Giacoumidis is the principal investigator of the world's-first high-capacity intensity-modulated and direct-detected optical Fast-OFDM system. He was the first to implement machine learning in optical communications (coherent optical OFDM) for fibre nonlinearity compensation.

Sofien Mhatli is a senior researcher at the Ecole Polytechnique de Tunis and a Telecommunications engineer at the Ministry of National Defense of Tunisia. His expertise is focused on designing and optimizing optical and microwave devices for applications in wireless telecom, sensing systems, and optical communication systems (SDH, WDM, PON, CO-OFDM, IM-DD...). Dr. Mhatli has deep knowledge in digital signal processing, algorithmic design (e.g. machine learning), and optical devices (e.g. MZM) for optical communications. The outcome of his research resulted in prestigious peer-reviewed journal papers and top conference presentations (e.g. OSA Optics Letters, OFC, IEEE Photonics Journal, ECOC).

Mark F. C. Stephens has worked in the field of optical communications for ~15 years since his PhD at Bristol University and subsequent employment in R&D at Nortel, Marconi, and Ericsson. He is currently researching parametric optical amplification as a potential replacement for the EDFA in communications systems as part of a 3-year EPSRC project.

Athanasios Tsokanos is a senior lecturer in Computer Science at the University of Hertfordshire, UK. He holds a PhD in optical networking. In the past, he has worked in the networking industry and his current research interests are optical networks, machine learning and network security.

Jinlong Wei is currently a senior researcher at Huawei Technologies Duesseldorf GmbH, Munich, Germany. He was a senior engineer and Marie Curie fellow in ADVA Optical Networking SE, Germany from Sept. 2014 to Aug. 2016. Prior to that, he worked at University of Cambridge, UK as a research associate for 3 years starting from Aug. 2011. He received a PhD degree in Electronic Engineering from the University of Wales, Bangor, UK in 2011 and worked there as a post-doc researcher afterwards. He has participated in about 10 European national and international projects and realized a number of world-first system demonstrations with his colleagues, such as a real time 40 Gb/s lane rate PAM-4 based NG-PON and a real time 400 Gb/s inter-data center connection optical link. The results have been reported by leading media such as BBC, Reuters, BussinessWire etc. He has been contributing to the next generation 100 Gigabit Ethernet study within IEEE802.3. He has authored or co-authored more than 100 international journal/conference papers, including several invited and holds several US/Europe patents. He has served as a session organizer or a technical committee member of several international conferences.

Mary E. McCarthy biography not available at the time of submission.

Nick J. Doran has over 30-years research experience in high-speed and long distance optical communications. He led a research team at BT for 10 years on both theoretical and experimental investigations in ultra-high speed optical systems from 1981 to 1991. He jointly established the photonics research group at Aston University 1991-2000 specializing in soliton communication and processing. During this time the research was extensively funded by EPSRC and supported by industrial contracts with Marconi and KDD. In 2000 he established a start-up development within Marconi (SOLSTIS) to develop an ultralong communication system based on his research. In 2005 he took on the role of Head and Director of the Institute of Advanced Telecommunication (IAT) at Swansea University. He returned to Aston University in November 2013 and there runs two key research projects on nonlinear fibre amplification (EPSRC) Wideband Optical Communication Systems Using Phase-Sensitive/Insensitive Fibre Optical Parametric Amplifiers) and optical networks (FP7 DISCUS). Prof Doran has published over 200 papers and 20 patents on optical transmission and processing. He invented the concept of dispersion managed solitons and the extensively used Nonlinear Optical Loop Mirror (NOLM). He has a current H–index of 48 with over 8300 citations of his publications.

Andrew D. Ellis has previously worked for British Telecom Research Laboratories as a senior research engineer investigating the use of optical amplifiers and advanced modulation formats in optical networks and the Corning Research Centre as a senior research fellow where he led activities in optical component characterization. From 2003, he headed the Transmission and Sensors Group at the Tyndall National Institute in Cork, Ireland, where he was also a member of the Department of Physics, University College Cork and his research interests included the evolution of core and metro networks, and the application of photonics to sensing. He is now Professor of Optical Communications at Aston University where he is also deputy director of the Institute of Photonics Technologies (AIPT). He has published over 170 journal papers and over 25 patents in the field of photonics, primarily targeted at increasing capacity, reach and functionality in the optical layer. Prof. Ellis is a member of the Institute of Physics and the Institute of Engineering Technology, and is a Chartered Physicist. He is an associate editor of the journal Optics Express. Prof Ellis is a member of the Technical Program Committee of ECOC, chairing sub-committee 3 devoted to digital and optical signal processing in 2014.