Optics Letters

Comparison of DSP-based nonlinear equalizers for intra-channel nonlinearity compensation in coherent optical OFDM

Elias Giacoumidis,[1] Sofien Mhatli,[2] Tu Nguyen,[3] Son T. Le,[4] Ivan Aldaya,[5] Mary E. McCarthy,[4] Andrew D. Ellis,[4] and Benjamin J. Eggleton[1]

[1]Centre for Ultrahigh Bandwidth Devices for Optical Systems (CUDOS) and Institute of Photonics and Optical Science (IPOS), School of Physics, University of Sydney, Sydney, NSW 2006, Australia
[2]SERCOM-Lab, EPT Université de Carthage, La Marsa, Tunis 2078, Tunisia
[3]FPM-UMONS, 31 Dolez, B-7000 Mons, Belgium
[4]Aston Institute of Photonic Technologies (AIPT), Aston University, Birmingham, B4 7ET, England, UK
[5]Physics Institute, State University of Campinas, 777, Campinas, Brazil
*Corresponding author: e.giacoumidis@physics.usyd.edu.au

Received 2 February 2016; revised 27 April 2016; accepted 28 April 2016; posted 4 May 2016 (Doc. ID 258652); published 20 May 2016

A novel versatile digital signal processing (DSP)-based equalizer using support vector machine regression (SVR) is proposed for 16-quadrature amplitude modulated (16-QAM) coherent optical orthogonal frequency-division multiplexing (CO-OFDM) and experimentally compared to traditional DSP-based deterministic fiber-induced nonlinearity equalizers (NLEs), namely the full-field digital back-propagation (DBP) and the inverse Volterra series transfer function-based NLE (V-NLE). For a 40 Gb/s 16-QAM CO-OFDM at 2000 km, SVR-NLE extends the optimum launched optical power (LOP) by 4 dB compared to V-NLE by means of reduction of fiber nonlinearity. In comparison to full-field DBP at a LOP of 6 dBm, SVR-NLE outperforms by ∼1 dB in *Q*-factor. In addition, SVR-NLE is the most computational efficient DSP-NLE.

OCIS codes: (060.2330) Fiber optics communications; (060.4080) Modulation; (060.1660) Coherent communications.

http://dx.doi.org/10.1364/OL.41.002509

Endeavors to surpass the Kerr-induced nonlinearity limit have been performed by either inserting an optical phase conjugator (OPC) at the middle point of the link [1] or using electronic-based nonlinearity compensators (NLC) such as digital back-propagation (DBP) placed in the receiver [2] or transmitter [3], phase-conjugated twin-waves (PC-TW) [4], and nonlinear equalizers (NLEs) based on the inverse Volterra series transfer function (V-NLE) [5]. Unfortunately, the OPC significantly reduces the flexibility in an optically routed network requiring both symmetric second-order chromatic dispersion (CD) and power evolution; DBP is extremely complex, and PC-TW halves the transmission capacity. V-NLE has been considered as a simple and effective method for combating fiber nonlinearities; however, it still requires a significant amount of floating-point multiplications. Additionally, in coherent communication systems, the interaction between nonlinear phenomena, CD, and frequency fluctuations of source and local oscillators (LOs) results in stochastic nonlinear distortion, which can be partially mitigated using either frequency referenced carriers [3] or nonlinear mapping based on statistical learning such as artificial neural networks (ANN) [6] and support vector machines (SVM) [7].

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 and tolerance to CD and polarization-mode dispersion (PMD). However, due to its high peak-to-average power ratio, the nonlinear cross-talk effects among subcarriers such as cross-phase modulation (XPM) and four-wave mixing (FWM) are enhanced, causing a stochastic-like interference to the extent of becoming an insurmountable obstacle. Owing to the vulnerability of CO-OFDM in nonlinear distortion, it is envisaged that NLC will enhance the capacity and transmission-reach in coherent optical core networks [8], thus avoiding highly dissipative regeneration electronics [3]. However, NLC feasibility demands the employment of versatile (i.e., independent from link parameters) techniques of low complexity for real-time applications.

In this Letter, for the first time, V-NLE and full-field DBP-NLE are experimentally compared with a novel SVM-based regression (SVR) NLE in 40 Gb/s 16-quadrature amplitude modulation (16-QAM) CO-OFDM at 2000 km. In contrast to nonlinear classifiers such as ANN [6] and SVM [7], SVR projects the obtained data on a hyperplane where constellation regions are easier to decode. It is shown that SVR-NLE can extend the optimum launched optical power (LOP) by 4 dB compared to both linear equalization and V-NLE by means of reduction of fiber nonlinearity. In comparison to full-field

DBP-NLE at a LOP of 6 dBm, SVR-NLE outperforms by ~1 dB in Q-factor. In addition, it is shown that SVR is significantly less complex than full-field DBP and V-NLE.

Figure 1 depicts (a) the block diagram of the CO-OFDM receiver equipped with NLE, and (b) the proposed SVR-NLE comprised of k hidden nodes (support vectors) with each node being associated to each subcarrier k. The received symbols for each subcarrier $x\{k\}$ are processed by the NLE supported vectors which are scaled by weight values (i.e., the Lagrange multipliers) for each subcarrier $w_{k,i}$, after which the outputs for different k are summed.

Distribution of a noisy possible constellation point is learned during an initial training process. Once these distributions are learned, the detector can make a decision about new unknown observation symbols. The hyperplane is obtained through approximation of a nonlinear function using a set of kernels (sigmoid function) of l training dataset $\{(x^1, y^1), (x^2, y^2), \ldots, (x^l, y^l)\}$. In Fig. 2, an example is depicted for SVR showing how a dataset with noise can be extracted using the kernel "trick" and thus controlling "overfitting." In Fig. 2(a), the graphical representation depicts a data-set with noise using vectors of +1, −1 and, in Fig. 2(b), it is shown how the useful data could be extracted via a powerful kernel without being corrupted by the noisy data.

Afterward, SVR maps the data to a high-dimensional feature space using a nonlinear mapping φ (kernel-based sigmoid function); then linear regression is formulated by introducing the "ε-insensitive" loss function in the following form:

$$f(x, w) = \sum_{i=1}^{M} w_{k,i}\varphi_{k,i}(x) + b, \tag{1}$$

where $f(x, w)$ is the target linear model, $\varphi_{k,i}(x)$ denotes a set of nonlinear transformations of input x, and b is the bias term. The number of vectors in every hidden node is equal to the number of points of the constellation, i.e., M in (1) which, in the case of 16-QAM, is 16. Afterward, (1) can be learned

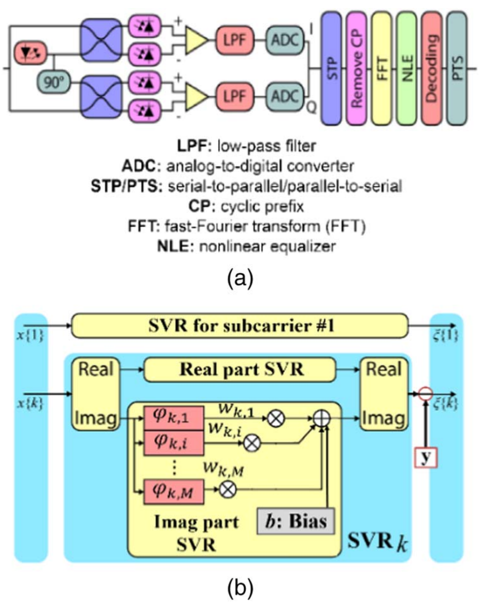

Fig. 1. (a) Block diagram of the CO-OFDM receiver equipped with NLE. (b) Proposed SVR-NLE.

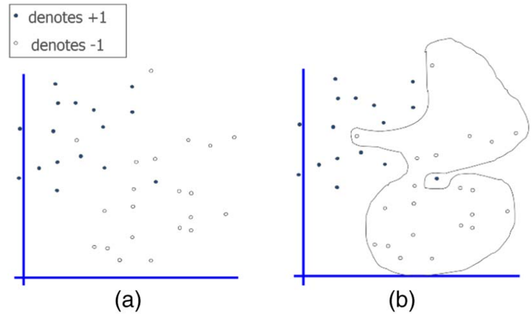

Fig. 2. SVR example: (a) dataset with noise and (b) example of using powerful kernel "trick" to distinguish useful data from noisy data.

through a training process by minimizing the error:

$$\psi(w, \xi) = \frac{1}{2}\|w\|^2 + C\sum(\xi_k^- + \xi_k^+), \tag{2}$$

where ξ_k^-, ξ_k^+ are slack variables [7] corresponding to the upper and lower bounds on the output function, and C is the penalty parameter that controls the trade-off between the slack variable penalty and the margin [7]. Depending on how much loss is ignored, the solution of (2) can be approximated by the Lagrange-based loss function $L(y, f(x, w))$ and, thus, the adopted "ε-insensitive" loss function can be expressed as

$$L_\varepsilon(y, f(x, \omega)) = \begin{cases} 0 & \text{if } |y - f(x, w)| \le \varepsilon \\ |y - f(x, w)| - \varepsilon, & \text{otherwise} \end{cases}. \tag{3}$$

The procedure of SVR involves two stages:

- Training:

(a) Arrange the data to form SVR packet with label (I and Q);
(b) Perform (I, Q) data scaling to [0,1];
(c) Select the sigmoid kernel function;
(d) Use cross-validation to find the best C and the standard single-mode fiber (SSMF)-induced nonlinearity parameter;
(e) Use C and nonlinearity parameter to build the SVR for the whole "training set."

- Testing:

(a) Approximate functions of the form presented in (1) with the "$\varepsilon -$ insensitive" loss function. Loss is zero if the difference between $f(x, w)$ and the measured value is $<\varepsilon$. Vapnik's "ε-insensitivity" loss function defines an ε tube around $f(x, w)$ [9]. If the predicted value is within the tube, the loss (error, cost) is zero while, for points outside, the loss is equal to the magnitude of the difference between the predicted value and the radius ε of the tube.
(b) Compare predicted labels (y-output) to pre-stored transmitted label for bit-error-rate (BER) estimation. In SVR, even if the processing of the initial training sequence is computational consuming, for a highly stable link, where CD and nonlinear effects do not change over time, the regression coefficients should only be found once. Moreover, there is no oversampling, as in V-NLE, because SVR-NLE is performed in a subcarrier-by-subcarrier OFDM process.

The block diagram of V-NLE is depicted in Fig. 3, which is similar to [5,6]. For V-NLE, third-order Volterra kernels were considered to reduce the complexity, which is identical to that reported in [5,6] to account for single-polarization 16-QAM CO-OFDM. In contrast to SVR-NLE, V-NLE is placed after the analogue-to-digital converters (ADCs), as depicted in Fig. 3, to relax digital signal processing (DSP) complexity by means of reducing the number of inverse fast Fourier transform (IFFT)/FFT blocks. V-NLE inherits some of the features of the hybrid time-and-frequency domain implementation such as nonfrequency aliasing and simple implementation. From Fig. 3, it can be clearly identified that CD, i.e., $(H_{CD})^k$, and the fiber nonlinearity are combated by the linear and nonlinear compensator tool, respectively.

Finally, 200 steps/span (denoted as full-field) were taken for DBP-NLE following a procedure similar to [2]. DBP-NLE is also placed after the ADCs in the receiver.

The number of floating-point real-valued multiplications required by SVR for decoding each OFDM symbol is

$$N_{\mathrm{SVR}} = 2 \cdot N_{\mathrm{SC}}(2^{N_{\mathrm{bits}}} + 1), \quad (4)$$

where N_{SC} is the number of subcarriers, and N_{bits} is the number of bits encoded in each subcarrier. The number of multiplications in full-field DBP is

$$N_{\mathrm{DBP}} = d_{\mathrm{link}}/d_{\mathrm{step}}[8N_{SC}K\log_2(N_{\mathrm{SC}}K) - 9KN_{\mathrm{SC}} + 16, \quad (5)$$

in which where d_{link} and d_{step} are the total link distance and the splitting step, respectively, and K is the oversampling factor. For a system under test with $N_{\mathrm{bits}} = 4$, $N_{\mathrm{SC}} = 512$ and $d_{\mathrm{link}} = 2000$ km and, assuming $K = 4$ and $d_{\mathrm{step}} = 1$ km, it is calculated that $N_{\mathrm{SVR}} = 17408$ whereas $N_{\mathrm{DBP}} = 145440000$, that is, a difference of ~4 orders of magnitude. On the other hand, the number of multiplications required by V-NLE is $N_{\mathrm{Volterra}} = (N_{\mathrm{span}} + 1)8N_{\mathrm{SC}}K\log_2(N_{\mathrm{SC}}K) + (20N_{\mathrm{span}} - 6)N_{\mathrm{SC}}K + 16(N_{\mathrm{span}} + 1)$ in which where N_{span} is the number of spans. The computational complexity of V-NLE depends on N_{span}, but not on N_{bits} while, as shown from (4), SVR-NLE does not depend on the link-related parameters, but on N_{bits}, since it is sensitive to the number of points in the constellation. Figure 4 shows a detailed quantitative comparison in terms of N_{SC} for different system parameters. For V-NLE, a K of 4 has been set, and the N_{span} has been varied (1, 10, 20); for DBP-NLE, a K of 4 has also been set, and the total link distance (d_{link}) has been varied (100 km, 1000 km, and 2000 km) whereas, for SVR-NLE, the N_{bits} has been swept ($k = 1, 2$, and 4). It is shown that, for all considered N_{SC} values, SVR-NLE outperforms both V-NLE and DBP-NLE, in terms of computational complexity. Even when comparing the best-case scenario of V-NLE, i.e., only 1 span, to the worst-case scenario of SVR i.e., 4 N_{bits}, the latter always outperforms. This difference increases according to the number of spans, which is the case of long-haul networks.

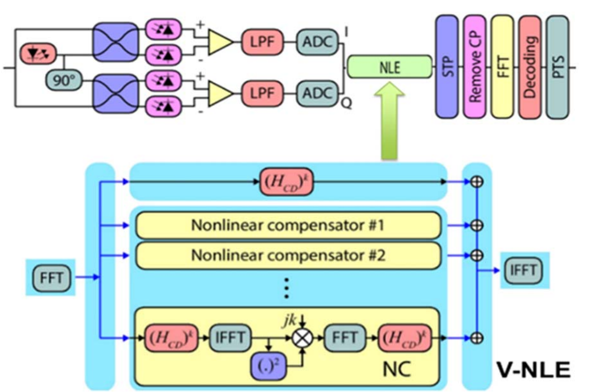

Fig. 3. Block diagram of the CO-OFDM receiver equipped with the benchmark V-NLE. LPF, low-pass filter; ADC, analogue-to-digital converter; STP/PTS, serial-to-parallel/parallel-to-serial; CP, cyclic prefix; (I)FFT, (inverse) fast-Fourier transform; NLE, nonlinear equalizer; NC, nonlinear compensator; H_{CD}, nonlinear system chromatic dispersion.

Figure 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 I-Q configuration. The DP-MZM was fed with OFDM I-Q components, which were generated offline. The transmission path at 1550.2 nm was a recirculating loop consisting of 20 × 100 km spans of Sterlite OH-LITE (E) fiber (attenuation, α, 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 first Erbium-doped fiber amplifier (EDFA), and a gain-flattening filter (GFF) was placed in the mid-stage of the third EDFA. The 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 LO. After downconversion, the baseband signal was sampled using a real-time oscilloscope operating at 80 GS/s and processed offline in MATLAB. Four-hundred OFDM symbols were generated using a 512-point IFFT; 210 subcarriers were modulated using 16-QAM, while the rest were set to zero. To eliminate the PMD-induced inter-symbol-interference, a cyclic prefix (CP) of 2% was included. The SVR training overhead was set at 10%, similar to [7], resulting in a training length of 40 symbols. The net bit rate was ~40 Gb/s. The offline OFDM demodulator included both timing synchronization and frequency offset compensation, as well as I-Q imbalance and CD compensation using an overlapped frequency domain equalizer employing the overlap-and-save method. The raw bit rate for all techniques, including linear equalization, was ~45.6 Gb/s. NLE was assessed by Q-factor measurements averaging over 10 recorded traces (~10^6 bits), which was

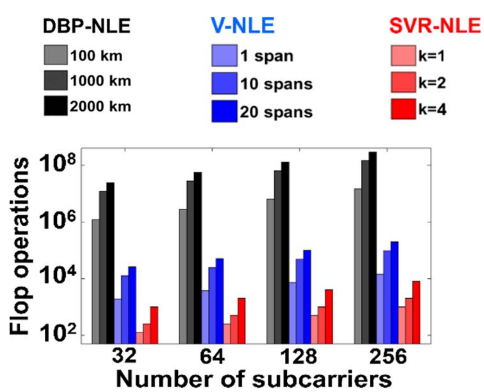

Fig. 4. Computational complexity comparison between SVR-, DBP-, and V-NLEs: the blue bars represent the computational complexity of V-NLE for a different subcarrier number, N_{SC}, and the number of spans, N_{span}; black-gray bars represent the computational complexity of DBP-NLE for a different subcarrier number, N_{SC}, and transmission lengths, whereas the red bars are for SVR-NLE considering various N_{SC}, and bits per subcarrier, N_{bits} (i.e., k).

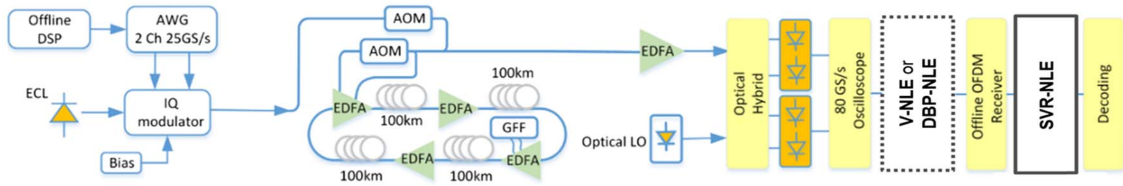

Fig. 5. Experimental setup of 40 Gb/s CO-OFDM equipped with NLE. ECL, external cavity laser; DSP, digital signal processing; AWG, arbitrary waveform generator; AOM, acousto-optic modulator; EDFA, Erbium-doped fiber amplifier; GFF, gain flatten filter; LO, local oscillator.

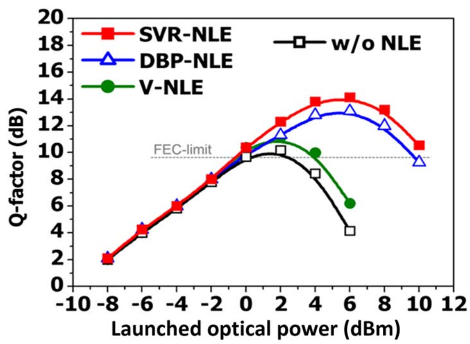

Fig. 6. *Q*-factor versus launched optical power (LOP) for SVR-NLE, DBP-NLE, and V-NLE and w/o NLE, for 40 Gb/s 16-QAM CO-OFDM at 2000 km of transmission.

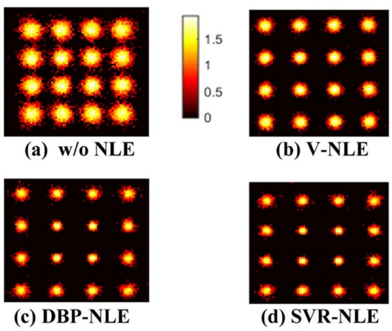

Fig. 7. Received 16-QAM constellation diagram of 40 Gb/s CO-OFDM for a LOP of 4 dBm: (a) w/o NLE [*Q*-factor = 8.2 dB], (b) with V-NLE [*Q*-factor = 10.4 dB], (c) with DBP-NLE [*Q*-factor = 12.5 dB], and (d) with SVR-NLE [*Q*-factor = 14 dB]. Inset: color bar.

estimated from the BER obtained by error counting after hard-decision decoding. The *Q*-factor is related to BER by $Q = 20\log_{10}[\sqrt{2}\,\mathrm{erfc}^{-1}(2\mathrm{BER})]$. For 16-QAM, a BER of 10^{-3} (FEC-limit) results in a *Q*-factor of ~9.8 dB.

In Fig. 6, the *Q*-factor against the LOP is plotted for 40 Gb/s CO-OFDM at 2000 km for SVR-NLE, DBP-NLE, and V-NLE and without (w/o) employing NLE. It is shown that SVR-NLE can extend the optimum LOP by 4 dB compared to both linear equalization and V-NLE by means of reduction of fiber nonlinearity. In comparison to full-field DBP-NLE at a LOP of 6 dBm, SVR-NLE outperforms by ~1 dB in *Q*-factor. This occurs due to the inability of both V-NLE and DBP-NLE to combat the stochastic-induced inter-subcarrier intermixing-based effects of FWM and XPM.

In Fig. 7, a comparison is depicted between the received 16-QAM constellations of 40 Gb/s CO-OFDM at a LOP of 4 dBm for all three aforementioned DSP-NLEs and w/o NLE. From these constellations it is evident that SVR-NLE condenses the constellation points more effectively in contrast to the two benchmark deterministic NLEs (i.e., V-NLE and DBP-NLE) by decreasing the symbol dispersion due to its stochastic-induced nonlinearity mitigation capability.

In conclusion, a novel SVR-based NLE was experimentally compared to V-NLE and full-field DBP-NLE in 40-Gb/s 16-QAM CO-OFDM at 2000 km. SVR-NLE extended the optimum LOP by 4 dB compared to V-NLE and outperformed by ~1 dB in *Q*-factor to full-field DBP-NLE at a LOP of 6 dBm. Moreover, SVR-NLE was significantly less complex than both full-field DBP-NLE and V-NLE.

Funding. Centre of Excellence (CE110001018); Laureate Fellowship (FL120100029); Engineering and Physical Sciences Research Council (EPSRC) (EP/J017582/1, EP/L000091/1); FAPESP (2015/04113-0).

REFERENCES

1. I. D. Phillips, M. Tan, M. F. C. Stephens, M. E. McCarthy, E. Giacoumidis, S. Sygletos, P. Rosa, S. Fabbri, S. T. Le, T. Kanesan, S. K. Turitsyn, N. J. Doran, P. Harper, and A. D. Ellis, in *Optical Fiber Communications Conference and Exhibition (OFC)*, San Francisco, CA, (2014), paper M3C.1.
2. G. Gao, J. Zhang, and W. Gu, IEEE Photon. Technol. Lett. **25**, 717 (2013).
3. E. Temprana, E. Myslivets, B. P.-P. Kuo, L. Liu, V. Ataie, N. Alic, and S. Radic, Science **348**, 1445 (2015).
4. X. Liu, A. R. Chraplyvy, P. J. Winzer, R. W. Tkach, and S. Chandrasekhar, Nat. Photonics **7**, 560 (2013).
5. E. Giacoumidis, I. Aldaya, M. A. Jarajreh, A. Tsokanos, S. T. Le, F. Farjady, A. D. Ellis, and N. J. Doran, IEEE Photon. Technol. Lett. **26**, 1383 (2014).
6. M. A. Jarajreh, E. Giacoumidis, I. Aldaya, S. T. Le, A. Tsokanos, Z. Ghassemlooy, and N. J. Doran, IEEE Photon. Technol. Lett. **27**, 387 (2015).
7. M. Li, S. Yu, J. Yang, Z. Chen, Y. Han, and W. Gu, IEEE Photon. J. **5**, 7800312 (2013).
8. S. T. Le, M. E. McCarthy, N. Mac-Suibhne, M. Al-Khateeb, E. Giacoumidis, N. J. Doran, A. D. Ellis, and S. K. Turitsyn, J. Lightwave Technol. **33**, 1308 (2015).
9. V. Vapnik, *Secaucus* (Springer, 1982).