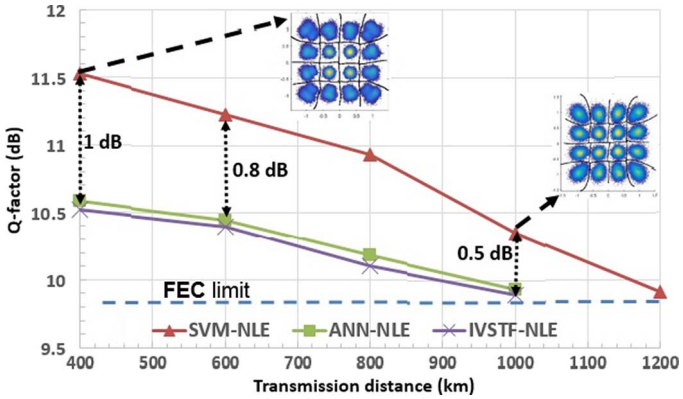
IEEE Photonics Journal
An IEEE Photonics Society Publication

Open Access

Fiber Nonlinearity Equalizer Based on Support Vector Classification for Coherent Optical OFDM

Volume 8, Number 2, April 2016

Tu Nguyen
Sofien Mhatli
Elias Giacoumidis
Ludo Van Compernolle
Marc Wuilpart
Patrice Mégret



DOI: 10.1109/JPHOT.2016.2528886

IEEE photonics SOCIETY

IEEE

Fiber Nonlinearity Equalizer Based on Support Vector Classification for Coherent Optical OFDM

Tu Nguyen,[1,4] Sofien Mhatli,[2] Elias Giacoumidis,[3] Ludo Van Compernolle,[4] Marc Wuilpart,[1] and Patrice Mégret[1]

[1]Faculté Polytechnique—Université de Mons, 7000 Mons, Belgium
[2]SERCOM-Laboratory, EPT Université de Carthage, Tunis 2078, Tunisia
[3]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
[4]Proximus SA, 1030 Bruxelles, Belgium

DOI: 10.1109/JPHOT.2016.2528886

Manuscript received January 9, 2016; revised February 6, 2016; accepted February 9, 2016. Date of publication February 11, 2016; date of current version February 23, 2016. This work was supported by the EU project ICONE, under Grant 608099, in collaboration with Proximus SA, Belgium. The work of E. Giacoumidis was supported by the Australian Research Council (CUDOS/IPOS, CE110001018). Corresponding author: T. Nguyen (e-mail: tuthanh.nguyen@umons.ac.be).

Abstract: A support vector machine (SVM)-based classification nonlinear equalizer (NLE) is demonstrated, for the first time, in coherent optical orthogonal frequency-division multiplexing (CO-OFDM). For a 40-Gb/s 16 quadrature amplitude modulated (16QAM) CO-OFDM system at 400 km of transmission, SVM-NLE reduces the fiber-induced nonlinearity penalty by about 1 dB in comparison to the benchmark artificial-neural-network (ANN)-based and inverse-Volterra-series-transfer-function-based NLEs.

Index Terms: Support vector machine (SVM), fiber nonlinearity equalizer, artificial neural network (ANN), inverse Volterra series transfer function.

1. Introduction

Thanks to the rapid development of digital signal processing (DSP) in cooperation with coherent reception, the fiber impairments could be compensated in electrical and digital domains. This reduces the cost of the network infrastructure and enables the feasibility of highly robust transceivers. For the application of high speed, long-haul, and multi-wavelength optical networks, the coherent optical orthogonal frequency division multiplexing (CO-OFDM) technique has been considered as a potential candidate since it brings many advantages over single carrier modulation such as the capability of inter-symbol interference mitigation using simple DSP and higher bandwidth efficiency [1], [2]. One major drawback of OFDM is its high peak-to-average power ratio (PAPR), which leads to well-known nonlinear distortion in electrical power amplifiers and digital-to-analog converters (DACs) for broadband copper wire and wireless communication systems. In optical communications, the fiber channel introduces Kerr-induced nonlinear phase noise (i.e. self-phase modulation (SPM)) and nonlinear crosstalk effects among sub-carriers such as cross-phase modulation (XPM) and even four-wave mixing (FWM), which are considerably affected by the high PAPR. Obviously, higher launched power is necessary to maintain a

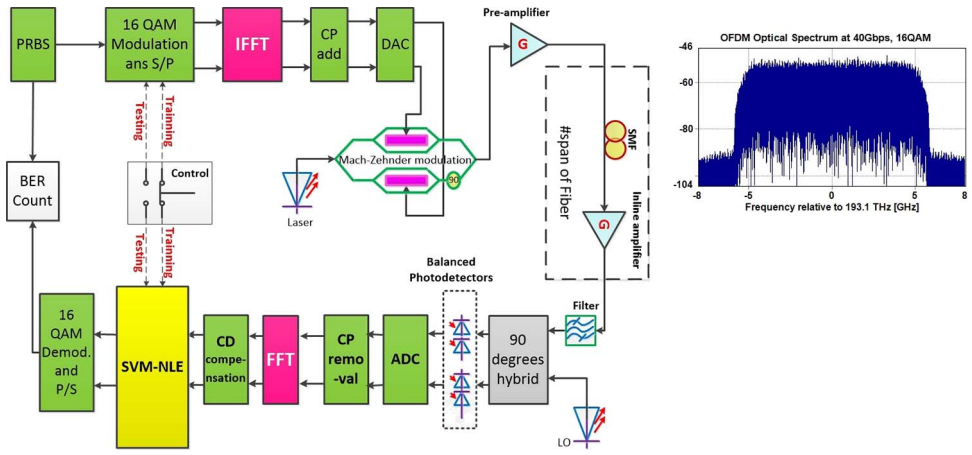


Fig. 1. Block diagram of 40 Gps 16QAM single-polarization CO-OFDM equipped with SVM-NLE. (Inset) Launched optical CO-OFDM spectrum.

sufficient signal-to-noise ratio (SNR) level when multi-level quadrature amplitude modulation (QAM) is employed in order to enhance spectral efficiency. However, high launched optical power results to enhanced nonlinear distortion due to the Kerr-effect which is proportional to the intensity of the power, thus producing more distortions in the constellation diagram of a multi-level modulation format (e.g., 16QAM)).

Numerous DSP techniques have been proposed for nonlinearity mitigation in single channel OFDM and wavelength division multiplexing (WDM)-based OFDM systems such as PAPR reduction techniques [2]–[4] and digital back propagation (DBP) [5]. In general, the main disadvantages of these techniques are the extensive use of fast Fourier transform (FFT) and/or inverse FFT (IFFT) modules and the regular pilots requirements. Recently, inverse Volterra series transfer function (IVSTF) and artificial neural network (ANN) based nonlinear equalizations (NLE) have gained huge attention in CO-OFDM systems [6]–[9]. However, both IVSTF and ANN-NLE need prior fiber's information which is not easy to obtain. SVM has been theoretically implemented in [10] for single-carrier 16QAM where the impact of chromatic dispersion (CD) on the nonlinear interaction has been ignored.

In this paper, we numerically investigate, for the first time, the performance of SVM-NLE in 40 Gbps 16QAM CO-OFDM, taking into account all transmission impairments without any prior information of the channel by only employing a very small amount of OFDM training symbols. Moreover, we show that SVM-NLE in 40 Gbps CO-OFDM can reduce the nonlinearity penalty by about 1 dB at 400 km of transmission compared to the benchmark ANN-NLE and IVSTF-NLE. However, for transmission distances longer than 400 km and up to 1200 km, the performance improvement by SVM-NLE is less than 1 dB due to the enhanced nonlinear inter-sub-carrier intermixing effects.

2. Support Vector Machine Based Nonlinearity Equalizer (SVM-NLE) in CO-OFDM

2.1. Description of CO-OFDM System

The CO-OFDM was first proposed in [11], and we adopted a similar design in our simulation analysis. Fig. 1 shows the block diagram of 16QAM CO-OFDM system equipped with SVM-NLE. The long-haul optical link consists of multiple spans of 100 km standard single mode fiber (SSMF) in which Erbium-doped fiber amplifiers (EDFA) were used to compensate fiber attenuation. Moreover, each EDFA adds amplified spontaneous emission (ASE) noise which is considered as an additive white Gaussian noise process.

The source bits from pseudo-random binary sequence (PRBS) module are fed into a digital modulator to form complex quadrature amplitude symbols. Via serial-to-parallel module, this high-stream of symbols is then re-arranged into multiple low rate streams of symbols where each low-speed stream of symbols corresponds to one input of multi-carrier modulator (i.e., IFFT). After IFFT, the baseband OFDM signal in one symbol period can be expressed as [1], [2]

$$s(n) = \sum_{k=0}^{N_{sc}-1} c_k e^{\frac{j2\pi kn}{N_{sc}}}, \tag{1}$$

where N_{sc} is the number of sub-carriers, and c_k is the kth QAM symbol. For simplicity, we assume that the pulse shape is ideal. The continuous complex baseband envelope of OFDM in one period therefore can be written as

$$s(t) = \sum_{k=0}^{N_{sc}-1} c_k e^{j2\pi f_k t} \Pi(t), \tag{2}$$

where $f_k = k/T_s$ is the frequency of the kth sub-carrier, with T_s being the useful OFDM symbol period. A cyclic prefix (CP) is added to create the complete OFDM signal with total symbol period of $T_{\text{OFDM}} = T_s + T_{cp}$. The ideal pulse shape, $\Pi(t)$, is given by

$$\Pi(t) = \begin{cases} 1, & \text{if } 0 < t < T_{\text{OFDM}} \\ 0, & \text{other.} \end{cases} \tag{3}$$

In-phase (I) and quadrature (Q) parts of OFDM signal are used to drive two Mach–Zehnder modulators (MZM). In order to obtain the best linear conversion between the OFDM signal and optical field and eliminate chirp phenomenon [1], [12], each MZM is biased at minimum transmission point and operated in push-pull configuration. The output optical signal after MZM is

$$E(t) = \sqrt{P_s} e^{j(\omega_s t + \phi_s)} \sum_{k=0}^{N_{sc}-1} c_k e^{j2\pi f_k t} \Pi(t), \tag{4}$$

where P_s, ω_s, and ϕ_s are average power, angular frequency, and phase of laser at transmitter, respectively. This OFDM optical signal is then launched into the fiber and propagated to the receiver side. At the receiver, the optical signal after N_{span} consecutive amplifier stages is detected and converted into electrical domain by 90^0 hybrid coherent photodetectors. The carrier phase noise due to imperfect laser linewidth is out of the scope of this paper as we want to isolate the fiber impairments from other source of noises. The complex photo-current after the two pairs of photodetectors consists of both I and Q components and is given by [2], [11]

$$I(t) = I_I(t) + jI_Q(t) = 2E_s E^*_{\text{LO}}, \tag{5}$$

where $I_I(t)$ and $I_Q(t)$ are the output currents after each pair of photodetectors, respectively. E_s and $E_{\text{LO}} = \sqrt{P_{\text{LO}}} e^{j(\omega_{\text{LO}} t + \phi_{\text{LO}})}$ are the optical incoming signal and optical local oscillator (LO), respectively. P_{LO}, ω_{LO}, and ϕ_{LO} are power, angular frequency and initial phase of LO. Due to fiber impairments and ASE noise, the optical signal at the receiver E_s can be expressed as

$$E_s(t) = \sqrt{P_s} e^{j(\omega_s t + \phi_s)} e^{j\varphi(t)} \Pi(t) \sum_{k=0}^{N_{sc}-1} c_k e^{j2\pi f_k t} e^{j\theta(k)} + w(t), \tag{6}$$

where $\varphi(t) = \gamma \|E(t))\|^2 N_{\text{span}} L_{\text{eff}}$ is the nonlinear phase noise term, with $\gamma = 2\pi n_2/(A_{\text{eff}}\lambda_0)$ being the nonlinear coefficient. $L_{\text{eff}} = (1 - e^{-\alpha L_{\text{SMF}}})/\alpha$ and A_{eff} are the effective length and area of SSMF, respectively. n_2 is the nonlinear reflective index of fiber and λ_0 is the center wavelength. The sign $\|.\|$ is the norm operator. This nonlinear phase noise expression is clearly proportional to the instantaneous power of the received signal. $w(t)$ is the accumulated ASE noise which is

generated by EDFAs and is given by $w(t) \sim \mathcal{N}(0, N_{\text{span}}\sigma^2_{\text{ASE}})$. A 80 GHz optical band-pass filter centering at 193.1 THz is placed just before the coherent receiver to limit the bandwidth of ASE noise. Linear phase shift on each OFDM sub-carrier due to CD is expressed by $\theta(k)$ term in (6). The mathematical evaluation of linear phase shift is [1], [11]

$$\theta(k) = \frac{\pi D_t c f_k^2}{f_0^2}, \tag{7}$$

where $D_t, c, f_k,$ and f_0 are dispersion coefficient, speed of light in free space, frequency of kth sub-carrier, and center frequency of optical carrier $(f_0 = \lambda_0/c)$, respectively. Obviously, the linear phase shifts can be easily compensated by linear filtering.

By assuming perfect coherent reception, i.e. $\omega_{\text{LO}} = \omega_s$, standard resistance of 1 Ω, power of LO at $1W$, and initial phases of transmitter laser and LO at both 0^0, we derived the received complex OFDM envelope, $r(t)$, by substituting (6) into (5)

$$r(t) = I(t) = 2\sqrt{P_s}e^{j\varphi(t)}\Pi(t)\sum_{k=0}^{N_{sc}-1} c_k e^{j2\pi f_k t}e^{j\theta(k)} + w(t). \tag{8}$$

After analog-to-digital converter (ADC), the CP is first removed from received complex baseband signal, $r(t)$. The time domain OFDM signal is then converted into frequency domain by the FFT. In the frequency domain, CD (the linear phase noise term $\theta(k)$ in (8)) will be compensated by a simple filter following the inverse function of (7). After CD compensation, the SVM-NLE is used to learn the nonlinear distortion (as expressed by the nonlinear term $(\varphi(t))$ in (8)), which is a result of fiber's nonlinear impairments such as SPM and inter-sub-carrier XPM and FWM. After SVM-NLE, nonlinear boundaries are formed and the detector uses this information to perform demodulation process. The output bits are finally converted into a serial stream of bits and compared to the transmitted bits to estimate bit-error-rate (BER). SVM-NLE procedure will be discussed in detail in the following section.

2.2. Support Vector Machine: Introduction and Procedure of Equalization

Our SVM model follows the procedure of Vapnik in [13]. The main concept is that SVM is used as a two-class classifier. By combining multiple two-class SVMs, a multi-class classifier model can be built. For example, the 16QAM signal is a multi-class signal where each cluster in constellation diagram belongs to one class. Distribution of noisy possible constellation points is learnt by training process. Let us take a two-class classifier for illustration: The hyperplane is obtained through training process from N pairs of training vectors $\{x_k, y_k\}, k = 1, \ldots, N$, where x_k and y_k are input and output patterns, respectively. In binary classification, each training vector corresponds to a label $y_k \in \{1, -1\}$. The objective of SVM is constructing a classifier $f(x)$ based on the training data [13], [14]

$$\begin{aligned} f(x) &= \text{sign}\left[\left(w^T\Phi(x_i) + b\right)\right] \\ &= \text{sign}\left[\sum_{k=1}^{N}\alpha_k y_k K(x, x_k) + b\right] \end{aligned} \tag{9}$$

where w, α_k, and b are weight, support vectors and bias terms which are determined via training process. The mapping function, $\Phi(.)$, is used to transfer the training data, x_k, into a higher dimensional space in which they might be linearly separated. In practice, Kernel "trick" helps for nonlinear decision on mapping with low computational complexity. The Kernel "trick" is defined as $K(x_i, x_j) \equiv \Phi(x_i)^T\Phi(x_j)$. There are some useful Kernels which have been proposed by researchers worldwide such as linear, polynomial and radial basis function (RBF) [14]. In this paper, we employ RBF Kernel function in which only dot products are computed: $K(x_i, x_j) \equiv \exp(-\gamma_{\text{SVM}}\|x_i - x_j\|^2), \gamma_{\text{SVM}} > 0$, with γ_{SVM} being the Kernel parameter.

In multi-class nonlinear noisy 16QAM constellation classification problem, we employed the one versus rest rule [14] in which a received constellation point would belong to a certain cluster if that cluster accepted a received symbol and other clusters rejected it. In order to increase the performance of SVM classification in noisy environment, least square SVM is introduced. The least square SVM provides solution to the optimization problem [14]

$$\min_{w,b,\xi}\left(\frac{1}{2}w^T w + C\sum_{i=1}^{N}\xi_i\right), \tag{10}$$

subject to constrain

$$y_i\left(w^T\Phi(x_i)+b\right) \geqslant 1-\xi_i, \xi_i \geqslant 0, \tag{11}$$

where ξ is the slack variable showing error term. C is the regularization parameter.

After CD compensation using training symbols, the I and Q components of complex symbols are fed into SVM-NLE. In this module, the SVM classifier is formed with the optimal decision boundaries (hyperplanes) which are used to classify in which cluster the received symbols should be allocated. The whole procedure of SVM-NLE consists of two stages which are listed in detail bellow.

1) Training:

- Arrange received complex symbols to format SVM packet (label, in-phase—I, and quadrature—Q).
- Scale I, Q to [0,1] and choose the RBF Kernel function.
- Perform cross-validation to find best C and γ_{SVM}.
- Use C and γ_{SVM} to train SVM detector.

2) Testing:

- Input testing symbols.
- Compare predicted labels (output of SVM-NLE) to pre-stored transmitted symbols to evaluate symbol error rate and BER after 16QAM demodulation.

3. Simulation Set-Up and Results

Numerical simulation of the CO-OFDM system equipped with SVM-NLE was performed by MATLAB/VPI TransmissionMarker co-simulation environment. The system operated at a net bit rate of 40 Gbps before CP extension. A transmission link of up to 12 spans with 100 km span-length was considered. We adopted a low number of 64 sub-carriers in order to reduce the PAPR in CO-OFDM. The CP length was 12.5%. In-line EDFA has a noise figure (NF) of 6 dB. For the SSMF, the following parameters were used: fiber loss $\alpha_{SMF} = 0.2$ dB km^{-1}, CD parameter $D = 17$ ps km^{-1}nm^{-1}, dispersion slope of 0.08 ps km^{-1}nm^{-2}, nonlinear Kerr coefficient of 2.6×10^{-20} m^{2}W^{-1}, and effective core area $A_{eff} = 80 \times 10^{-12}$ m^2.

For all simulations, we utilized 5000 training symbols for the training process in order to obtain optimal hyperplanes, which accounts for less than 5% of the total symbols. Generally speaking, the training process takes place once, at the beginning of the transmission. The SVM model was conducted at the receiver side by adopting the Library for SVM (LIBSVM)—an open source for machine learning which was developed by National Taiwan University [14]. LIBSVM implements the sequential minimal optimization algorithm for kernelized SVM, supporting both classification and regression. Moreover, the cross-validation with small grid step-size search was employed to obtain the optimal values of RBF's parameters.

System performance was evaluated by BER which was obtained via error counting and Q-factor. The relationship between Q-factor and BER is $Q(\text{dB}) = 20\log_{10}[\sqrt{2}\text{erfc}^{-1}(2\text{BER})]$. Common threshold before forward error correction (FEC) for 16QAM is 10^{-3}, which is equivalent in Q-factor at 9.8 dB.

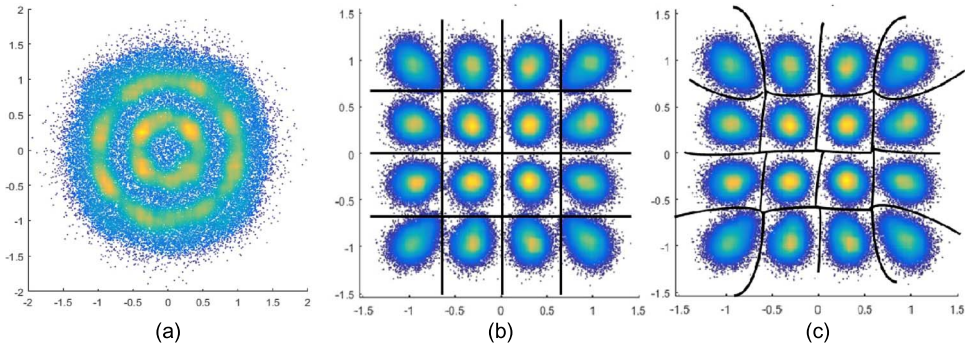


Fig. 2. Received constellation at the launched power of −6 dBm for a transmission distance of 1000 km. (a) Before CD compensation, (b) after CD compensation with hard decision boundaries, and (c) after SVM-NLE with nonlinear boundaries.

The transmitted optical OFDM spectrum is shown in inset of Fig. 1. Since we used the pulse shape Raised-cosine filter with roll-off factor of 0.2, the bandwidth of 40 Gbps 16QAM signal is approximately 12 GHz.

Fig. 2 shows the received constellation before CD compensation Fig. 2(a), after CD compensation Fig. 2(b) and after SVM-NLE Fig. 2(c) at −6 dBm of launched power for a transmission of 1000 km. It can be easily observed that the circular constellation which is caused by linear phase noise term on each sub-carrier can be equalized by a simple filter according to the inverse function of (7), and grouped tightly into 16 ideal constellation points as showed in Fig. 2(b). Residual phase noise on each constellation point after linear compensation is the result of ASE noise and nonlinear phase noise terms as illustrated in (8). Looking at the constellation diagram of the received symbols after CD compensation, we can observe that the impact of nonlinear phase noise depends on the instantaneous level of signal's strength as discussed in Section 2.1. As Fig. 2(b) reveals, the four inner constellation clusters are least affected by nonlinear phase noise whereas the outer rectangular constellation ring is more distorted. Consequently, normal hard decision thresholds of the 16QAM demodulator need to be adjusted nonlinearly according to the distribution of received symbols as illustrated in Fig. 2(c). After the end of the training process, these nonlinear boundaries are defined and used to classify the 16QAM symbols in the testing set.

In Fig. 3, we present Q-factor versus launched optical power for up to 1200 km of the 40 Gbps CO-OFDM system with and without SVM-NLE. It is shown that for each fiber distance, there is an optimum launched optical power where the best of Q-factor was obtained. The optimum launched power level is around −6 dBm, which is identical to that reported in [9] confirming the validity of the developed theoretical model. Moreover, it is shown that SVM-NLE can extend the maximum achievable transmission distance to 1200 km for the targeted FEC threshold. In comparison without (w/o) using NLE, the adopted SVM-NLE can reduce the intra-channel nonlinearities just around 0.5 dB at the optimum optical launched power. This improvement is a result of the nonlinear boundary decision of SVM classification. It is worth to mention that SVM-NLE has the ability of virtually correcting the skew in the received constellation diagram if training-assisted phase noise compensation is not adequate. However, it was realized that this skew could be easily corrected by a small amount of training symbols without the requirement of SVM, and hence, SVM-NLE was only required in CO-OFDM to reduce fiber-induced nonlinearities.

In Fig. 4, a performance comparison between SVM-NLE, ANN-NLE, and IVSTF-NLE is depicted at an optimum launched optical power of −6 dBm for the 40 Gbps 16QAM CO-OFDM system. The adopted theoretical models for the benchmark ANN and IVSTF NLEs are identical to [7], [9] and all simulation parameters were tuned accordingly. From Fig. 4, it is evident that SVM-NLE can improve the nonlinear tolerance in comparison to ANN-NLE and IVSTF-NLE by about 1 dB and 0.5 dB in Q-factor for transmission distances of 400 km and 1000 km,

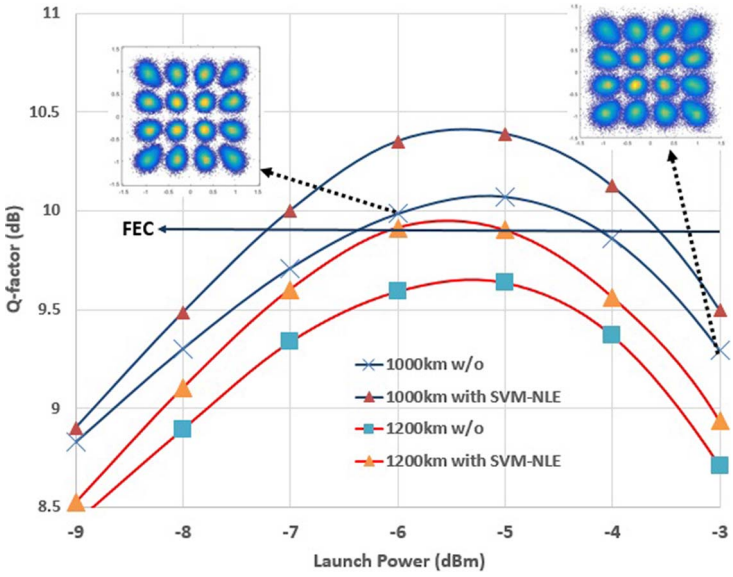


Fig. 3. Q-factor vs. launched optical power with and w/o SVM-NLE for 40 Gbps CO-OFDM at 1000 km and 1200 km, respectively.

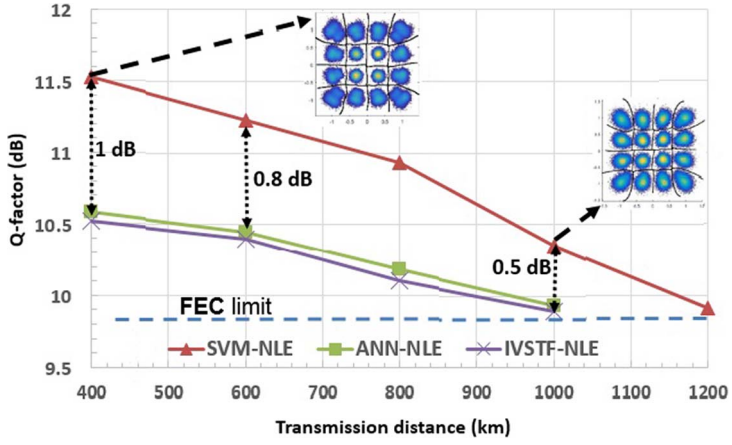


Fig. 4. Q-factor vs. distance of SVM-NLE, ANN-NLE, and IVSTF-NLE for 40 Gbps CO-OFDM at a launched optical power of –6 dBm.

respectively. When the transmission distance increases, the interaction between fiber's nonlinearities and the ASE noise is also enlarged considerably and, thus, the Q-factor improvement trend is declined, as showed in this figure. The performance improvement of SVM-NLE and ANN-NLE (slightly better) over deterministic NLE (i.e. IVSTF) in CO-OFDM can be explained by the inter-sub-carrier FWM and XPM products being so complicated that appear random, and hence, deterministic NLE cannot combat properly nonlinearities. Consequently, the introduction of NLEs that can combat stochastic impairments (i.e. SVM and ANN) is of great importance. On the other hand, SVM-NLE always outperforms ANN-NLE due to the fact that SVM can solve the problem of "overfitting" and local minima and that ANN uses two separate processors for real and imaginary part, thus ignoring cross information.

The number of floating-point real-valued operations required by SVM and ANN for decoding each OFDM symbol is given by $N = 2N_{sc}(2^{N_{bits}} + 1)$, where N_{bits} is the number of bits encoded on each subcarrier (i.e. $N_{bits} = 4$ for 16QAM). On the other hand, the number of operations

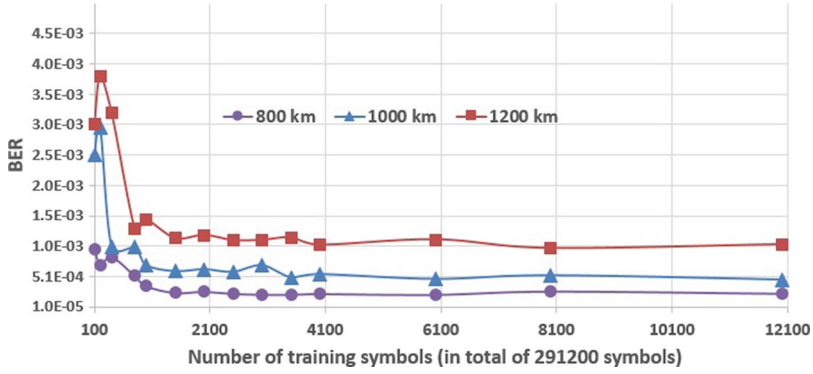


Fig. 5. Impact of training symbols number on performance of SVM-NLE.

required by IVSTF-NLE is given by $N_{\_\text{Volterra}} = (N_{\text{span}} + 1)8N_{sc}\log 2(N_{sc}K) + (20N_{\text{span}} - 6)N_{sc}K + 16(N_{\text{span}} + 1)$, where N_{span} is the number of spans, and K is the oversampling factor. The computational complexity of IVSTF-NLE depends on N_{span} but not on N_{bits}, while both SVM and ANN NLEs do not depend on the link-related parameters but on N_{bits} since they are sensitive to the number of points in the constellation. Clearly, for any type of modulation format and transmission-reach both SVM and ANN significantly reduce the computational complexity of the equalizer. SVM compared to ANN has similar computational complexity since they are both nonlinear classifiers providing similar sparsity to the system, and therefore, the amount of set of sparse supported vectors and artificial neurons is identical. However, in terms of performance SVM outperformed to ANN, as revealed from our results.

Finally, the impact of number of training symbols is shown in Fig. 5. Three scenarios were considered at a launched optical power of –6 dBm: 800 km, 1000 km, and 1200 km. It is shown that the SVM-NLE performance will be saturated fast, at around 1500 training symbols (< 1%) for all cases. As mentioned before, the training process takes place once, at the beginning of the transmission since the fiber could be seen as time-invariant channel and thus the impact of number of training on net data rate can be ignored in this study. In order to guarantee the sufficient optimum learning of SVM (around the FEC-limit) in all simulation scenarios, we adopted 5000 training symbols for all of our numerical validations, as discussed above.

4. Conclusion

We have implemented an SVM-based NLE, for the first time, in 40 Gbps 16QAM CO-OFDM to increase the nonlinear tolerance in long-haul transmission. We evaluated the performance of SVM-NLE in different launched power levels and fiber lengths by numerical simulation and compared it with the benchmark ANN-NLE and IVSTF-NLE. It was shown that SVM-NLE reduced the nonlinearity penalty by about 1 dB at 400 km of transmission compared to ANN-NLE and IVSTF-NLE.

Acknowledgment

The authors would like to thank Dr. I. Aldaya from the Physics Institute, State University of Campinas, Campinas, Brazil, for his expertise in improving the manuscript.

References

[1] J. Armstrong, "OFDM for optical communications," *J. Lightw. Technol.*, vol. 27, no. 3, pp. 189–204, Feb. 2009.
[2] H. Bao and W. Shieh, "Transmission simulation of coherent optical OFDM signals in WDM systems," *Opt. Exp.*, vol. 15, no. 8, pp. 4410–4418, Apr. 2007.
[3] B. Goebel, S. Hellerbrand, N. Haufe, and N. Hanik, "PAPR reduction techniques for coherent optical OFDM transmission," in *Proc. IEEE 11th ICTON*, Jun. 2009, pp. 1–4.

[4] J. Silva, A. Cartaxo, and M. Segatto, "A PAPR reduction technique based on a constant envelope OFDM approach for fiber nonlinearity mitigation in optical direct-detection systems," *IEEE/OSA Opt. Commun. Netw.*, vol. 4, no. 4, pp. 296–303, Apr. 2012.
[5] G. Gao, J. Zhang, and W. Gu, "Analytical evaluation of practical DBP-based intra-channel nonlinearity compensators," *IEEE Photon. Technol. Lett.*, vol. 25, no. 8, pp. 717–720, Apr. 2013.
[6] L. Liu *et al.,* "Intrachannel nonlinearity compensation by inverse Volterra series transfer function," *J. Lightw. Technol.*, vol. 30, no. 3, pp. 310–316, Feb. 2012.
[7] E. Giacoumidis *et al.,* "Volterra-based reconfigurable nonlinear equalizer for coherent OFDM," *IEEE Photon. Technol. Lett.*, vol. 26, no. 14, pp. 1383–1386, Jul. 2014.
[8] E. Giacoumidis *et al.,* "Fiber nonlinearity-induced penalty reduction in coherent optical OFDM by artificial neural network based nonlinear equalization," *OSA Opt. Lett.*, vol. 40, no. 21, pp. 1–4, Nov. 2015.
[9] M. Jarajreh *et al.,* "Artificial neural network nonlinear equalizer for coherent optical OFDM," *IEEE Photon. Technol. Lett.*, vol. 27, no. 4, pp. 387–390, Feb. 2015.
[10] Y. Han, S. Yu, M. Li, J. Yang, and W. Gu, "An SVM-based detection for coherent optical APSK systems with nonlinear phase noise," *IEEE Photon. J.*, vol. 6, no. 5, pp. 1–10, Oct. 2014.
[11] H. B. W. Shieh and Y. Tang, "Coherent optical OFDM: Theory and design," *Opt. Exp.*, vol. 16, no. 2, pp. 841–859, Jan. 2008.
[12] M. Seimetz, *High-Order Modulation for Optical Fiber Transmission*. Berlin, Germany: Springer-Verlag, 2009.
[13] V. Vapnik, *Estimation of Dependences Based on Empirical Data: Springer Series in Statistics*. Secaucus, NJ, USA: Springer-Verlag, 1982.
[14] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, May 2011.