





Effective handling of nonlinear distortions in CO-OFDM using affinity propagation clustering

JINLONG WEI,^{1,8}  ELIAS GIACOUMIDIS,^{2,9} JI ZHOU,³ MUTSAM A. JARAJREH,⁴ SHI LI,² IVAN ALDAYA,⁵  YANDONG LU,⁶ ATHANASIOS TSOKANOS,⁷ AND ANDRE RICHTER²

¹Huawei Technologies Düsseldorf GmbH, Riesstrasse 25, 80992 München, Germany

²VPIphotonics, Carnotstraße 6, 10587 Berlin, Germany

³Department of Electronic Engineering, College of Information Science and Technology, Jinan University, Guangzhou 510632, China

⁴Engineering Department, School of Computing, Fahad Bin Sultan University, Tabuk, United Arab Emirates

⁵Campus São Joao da Boa Vista, State University of São Paulo, 13876-750 São Paulo, Brazil

⁶Optical Access Network Hardware Development Dept, Huawei Technologies Co. Ltd, Dongguan, 523808, China

⁷University of Hertfordshire, Hatfield, Hertfordshire, UK

⁸jlongwei2@huawei.com

⁹elias.giacoumidis@vpiphotonics.com

Abstract: We experimentally demonstrate a system-agnostic and training-data-free nonlinearity compensator, using affinity propagation (AP) clustering in single- and multi-channel coherent optical OFDM (CO-OFDM) for up to 3200 km transmission. We show that AP outperforms benchmark deterministic and clustering algorithms by effectively tackling stochastic nonlinear distortions and inter-channel nonlinearities. AP offers up to almost 4 dB power margin extension over linear equalization in single-channel 16-quadrature amplitude-modulated CO-OFDM and a 1.4 dB increase in Q-factor over digital back-propagation in multi-channel quaternary phase-shift keying CO-OFDM. Simulated results indicate transparency to higher modulation format orders and better efficiency when a multi-carrier structure is considered.

© 2022 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

Increasing transmission speed in optical fiber communications is of utmost importance for many applications such as access, metropolitan networks, and long-haul networks due to the rise of the network traffic demand [1]. To increase data rates, one of the core difficulties is the optical Kerr effect, causing a variation in index of refraction which is proportional to the local irradiance of the light and is responsible for nonlinear optical effects such as cross-phase modulation (XPM) and four-wave mixing (FWM) that generates a 4th idler photon [1]. The optical Kerr effect is attributed to the so-called nonlinear Shannon capacity limit which sets an upper bound on the achievable data rate in optical fiber communications when using traditional linear transmission techniques [2]. On the other hand, there have been extensive efforts in attempting to approach the nonlinear Shannon limit through several fiber nonlinearity compensation techniques [1–6] that compensate exclusively deterministic Kerr-induced nonlinear effects. Albeit the Kerr-mediate nonlinear process is deterministic, it has been shown in [1] that the frequency uncertainty of many independent wavelength channels is transformed into time uncertainty through fiber transmission by chromatic dispersion (CD), making the nonlinear interaction appear random. This problem has been partially tackled by a combination of optical frequency combs (OFCs) [2] and digital back-propagation (DBP) [3]. Nevertheless, OFCs are impractical for flexible routed networks, while DBP is energy-ineffective and very complex for real-time signal processing [2,3].

On the other hand, recent popular coherent schemes introduce multiple carriers in a single frequency band to enhance spectral efficiency and flexibility such as coherent optical orthogonal frequency division multiplexing (CO-OFDM) [6] and Nyquist-wavelength division multiplexing (Nyquist-WDM). In these schemes, the resulting nonlinear interaction between subcarriers becomes quite complex making it appear more random rather than deterministic [7]. Recently, unsupervised and supervised machine learning such as K-means clustering [7] and artificial neural network regression [8] have been introduced to combat stochastic nonlinearities and transmitter imperfections (e.g., nonlinearity observed from a Mach-Zehnder modulator, MZM), performing blind soft-decision decoding (or in other terms nonlinear mapping) and non-blind nonlinear equalization (NLE), respectively. However, their performance benefit for multi-carrier schemes has been limited due to their inability to effectively compensate the strong nonlinear phase noise when using higher-order signal modulation formats on subcarriers.

In this work, we experimentally demonstrate the first system-agnostic (training-data-free) overlapping clustering algorithm using affinity propagation (AP) in single-channel 16-quadrature-amplitude-modulation (16-QAM) and WDM quaternary phase-shift keying (QPSK) CO-OFDM for long-haul transmission. The current work is a significant extension of the short-length conference paper presented in [9]. Results indicate that AP outperforms benchmark machine learning clustering algorithms in terms of signal quality (Q)-factor such as fuzzy-logic c-means (FL) and K-means. We also show that AP outperforms widely-adopted deterministic approaches, namely the full-step DBP (FS-DBP) and Volterra-based NLE [6], due to nonlinear inter-subcarrier mixing reduction induced from transmission and transmitter nonlinearities. Extensive numerical simulations indicate that AP is transparent to higher-order signal modulation formats and more efficient in a digital multi-carrier structure like OFDM when compared to single-carrier modulation.

2. Proposed algorithm and proof-of-concept by simulation

In machine learning clustering, defined data learn a set of centers such that the sum of squared errors between data points and their nearest centers is small [10]. When the centers are selected from actual data points, they are called “exemplars”. The well-known K-means clustering algorithm begins with an initial set of randomly selected exemplars and iteratively refines this set to decrease the sum of squared errors [7]. In AP, however, every symbol is a potential exemplar by viewing each symbol as a node that recursively transmits real-valued messages (separately for amplitude and phase) until a good set of exemplars and corresponding clusters emerge. “Messages” are updated by simple formulas that search for minima of an appropriately chosen energy function [10]. At any symbol in time, the magnitude of each message reflects the current affinity that one symbol has for selecting another symbol as its exemplar. Let for instance x_1 through x_n be a set of complex data (symbols), with no assumptions made about their internal structure, and let S be a function that quantifies the similarity between any two symbols, such that $S(x_i, x_j) > S(x_i, x_k)$; and considering x_i is more similar to x_j than to x_k . For this example, the negative squared Euclidean distance of 2 symbols is used i.e., for x_i and x_k . The diagonal of S (i.e. $S(i, i)$) is particularly important, as it represents the input preference, meaning how likely an input is to become an exemplar. When this is set to the same value for all inputs, it controls how many classes the algorithm can produce. An A value close to the minimum possible similarity produces fewer classes, however, a value close or larger to the maximum possible similarity produces many classes (initialized to the median similarity of all pairs of inputs). AP proceeds on how “appropriate” it would be for x_i to pick x_k as its exemplar, taking into account other points’ preferences. Responsibility R and availability A are initialized to zero and are viewed as log-probability tables, and afterward, AP is iteratively updated for R and A using the following expressions:

$$R(i, k) = s(i, k) - \max_{k' \neq k} \{a(i, k') + s(i, k')\} \quad (1)$$

$$A(i, k) = \min \left(0, r(k, k) + \sum_{i' \neq \{i, k\}} \max(0, r(i', k)) \right)_{\text{for } i \neq k} \quad (2)$$

The $a(i, k')$ in Eq. (1), is referred to the so-called “availability” matrix A , which contains the values i, k' that represent how “appropriate” it would be for x_i to pick $x_{k'}$ as its exemplar as mentioned above (taking into account other points’ preference for $x_{k'}$ as an exemplar). $r(i', k)$ is referred to the so-called “responsibility” matrix R with values i', k that quantify how well-suited $x_{i'}$ is to serve as the exemplar for x_i relative to other candidate exemplars. Note that for the availability matrix, a separate equation is used for updating the elements on its diagonal. The two matrices (R and A) essentially represent a graph where every data point is connected with all other points. Worth noting that the exemplars are extracted from the final updated matrices where ‘responsibility + availability’ is positive [10,11]. In the implementation process of the algorithm, note there is also a damping factor for numerical stabilization and can be regarded as a slowly converging learning rate with a value between 0.5 to 1. In Fig. 1, the steps of the AP algorithm are presented, while Fig. 2(a) shows the aforementioned AP iterative result of R and A for a simulated 20 Gb/s 4-QAM CO-OFDM single-channel with procedures identical to [7] (128 subcarriers) and over 3200 km of standard single-mode fiber (SSMF) transmission at the optimum launched optical power (LOP) of 1 dBm. At the digital receiver, after linear equalization to compensate phenomena such as CD, the AP clustering algorithm was added for processing separately I and Q to perform soft nonlinear decision boundaries on the linear equalized symbols. From Fig. 2(a), it is evident that 13 iterations are required for this signal to converge.

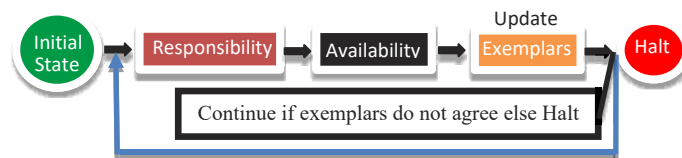


Fig. 1. Affinity propagation (AP) algorithm.

It is worth highlighting that while K-means and AP represent deterministic types of clustering, a “probabilistic type” also exists such as FL [7], permitting the symbols to fluctuate the data membership degree. In this work, we compare AP with FL and K-means clustering, which perform overlapping/soft and exclusive/hard clustering, respectively. Nevertheless, since AP also executes overlapping clustering, it is also considered as an overlapping deterministic clustering approach. The overlapping clustering ability of AP is demonstrated in Fig. 2(b), in which AP is compared with the exclusive (hard) clustering of K-means. For the sake of simplicity, we show the received 4-QAM constellation diagrams for the same simulated system demonstrated in Fig. 2(a). A Matlab/VPItransmissionMaker Optical Systems environment was used, in which digital modulation was implemented in Matlab and electro-optical components in VPItransmissionMaker Optical Systems. The number of symbols was chosen to be 2^{21} (random and unrepeated sequence). Results indicate that AP outperforms linear equalization (LE) (i.e., without using NLE) at optimum -3 dBm of LOP by 0.6 and 1 dB in Q-factor for 64- and 128-QAM, respectively.

It is worth noting that the benchmark K-means clustering is based on Lloyd’s algorithm [7], and is an iterative, data-partitioning algorithm that assigns n observations to exactly one of the K clusters defined by centroids. FL on the other hand, is a probabilistic clustering algorithm, permitting the symbols to fluctuate the data membership degree while being allocated into many clusters by minimizing an objective function. K-means and FL algorithms are detailed in [7].

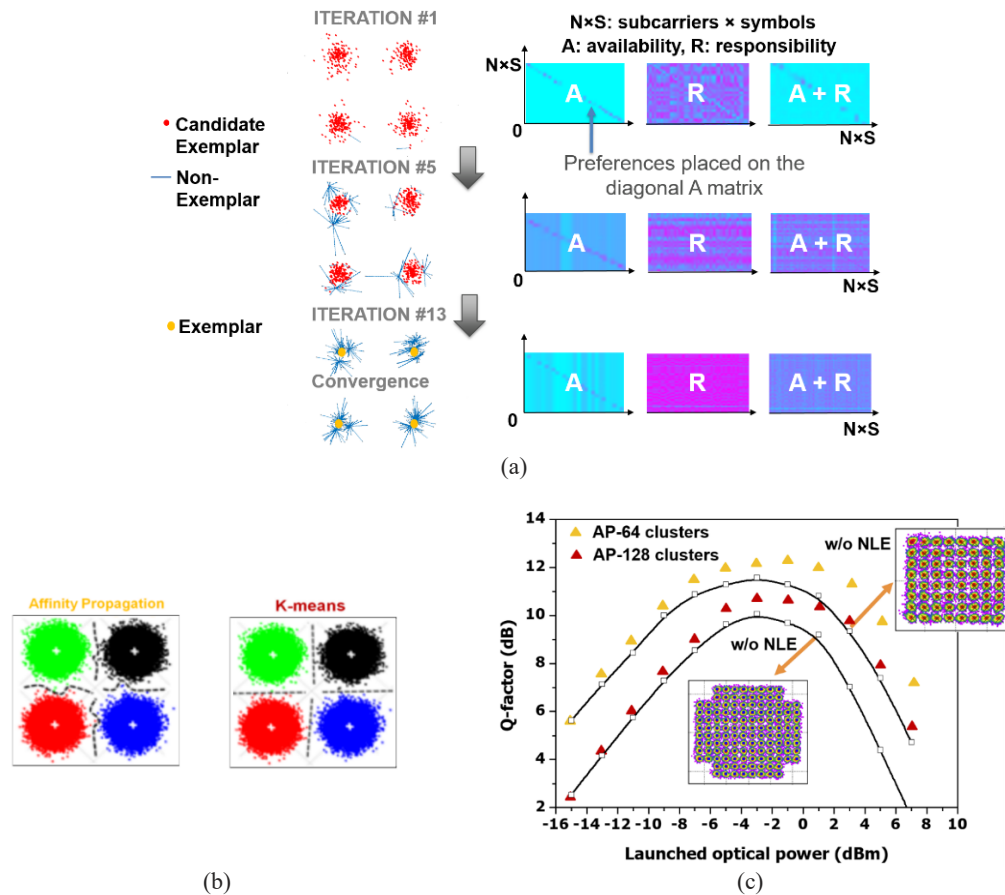


Fig. 2. (a) Example of AP algorithm steps. AP is tested for 4-QAM CO-OFDM. (b) Example of received constellation diagrams for AP overlapping clustering (left) and K-means exclusive clustering (right) for a simulated QPSK CO-OFDM system at 3200 km of SSMF. (c) AP simulated performance for 64- and 128-QAM CO-OFDM at 2000 km.

3. Experimental setup

Figure 3 depicts the experimental setup for single- and multi-channel CO-OFDM, where external cavity lasers of 100 kHz linewidth emitting at 193.4 THz were modulated using a dual-parallel MZM fed with 'offline' OFDM IQ components. The transmission link was established by a recirculating loop consisting of 20×100 km (single-channel) and 32×100 km spans (multi-channel) of OH-LITE fiber (attenuation of 18.9–19.5 dB/100 km) being switched by an acousto-optic modulator. The loop switch was located before the 1st Erbium-doped fiber amplifier (EDFA) and a gain-flattening filter was placed after the 3rd EDFA for both configurations to flatten the gain across the wavelengths of interest. For WDM-CO-OFDM, the transmitter was constituted of five distributed feedback lasers on a 100 GHz grid located between 193.5–193.9 THz, connected with a polarization-maintaining multiplexer. Using an amplified spontaneous emission (ASE) source, another 20 'dummy' channels of 10 GHz bandwidth were generated with a channel spacing of ~ 100 GHz. These channels covered 2.5 THz of bandwidth as depicted in the inset of Fig. 3. The optimum LOP was swept by controlling the output power of the EDFAs. At the receiver, the incoming signal was combined with a 100 kHz linewidth local oscillator for both single- and multi-channel configurations. After down-conversion, the 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 inverse fast Fourier transform (IFFT), 210 middle subcarriers were modulated using 16-QAM while the rest were set to zero. A cyclic prefix of 2% was included to eliminate inter-symbol interference. The OFDM demodulator for non-blind LE/Volterra-based NLE (V-NLE) included timing synchronization, IQ imbalance, CD, and frequency offset compensation [7] resulting in a net bit rate of ~ 20 and ~ 40 -Gb/s for QPSK and 16-QAM CO-OFDM, respectively. All NLEs were assessed by Q-factor measurements (related to bit-error-rate, BER, using $Q = 20 \log_{10}[\sqrt{2} \operatorname{erfc}^{-1}(2BER)]$) averaging over 10 recorded traces (2^{28} bits), which was estimated from the BER obtained by error counting after hard-decision decoding.

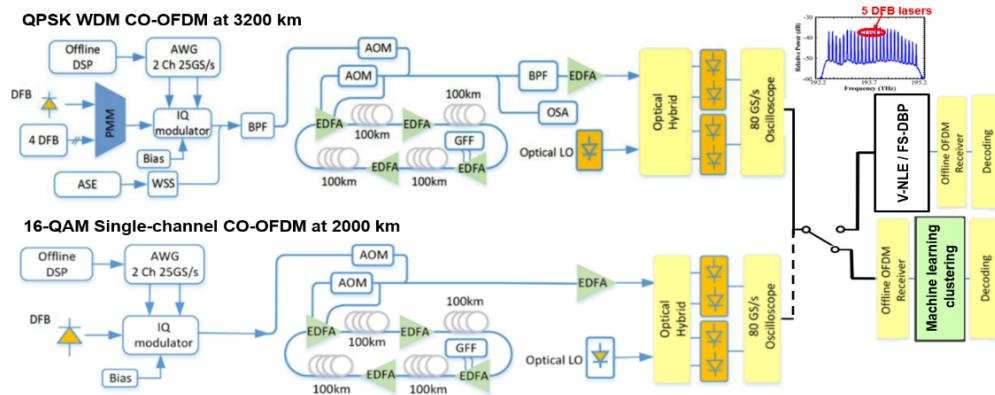


Fig. 3. Experimental setup for single- and multi-channel CO-OFDM configuration. Inset: Received spectrum for the multi-channel system. QPSK: quaternary phase-shift keying; WDM: wavelength division multiplexing; OFDM: orthogonal frequency division multiplexing; CO-OFDM: coherent optical OFDM; QAM: quadrature amplitude modulation; DFB: distributed feedback laser; DSP: digital signal processing; ASE: amplified spontaneous emission; PMM: polarization maintaining multiplexer; WSS: wavelength selective switch; AWG: arbitrary waveform generator; BPF: band-pass filter (optical); AOM: acousto-optic modulator; EDFA: Erbium-doped fiber amplifier; GFF: gain-flattening filter; OSA: optical spectrum analyzer; LO: local oscillator; V-NLE: Volterra-based nonlinear equalizer; FS-DBP: full-step digital back-propagation.

4. Experimental results

As shown in Fig. 4(a) for single-channel 16-QAM, AP offers a remarkable power margin extension of 4 dB compared to LE. Compared to all algorithms under test (i.e., FL, K-means, FS-DBP [40 steps/span], and V-NLE), the proposed algorithm outperforms over the entire range of LOPs. At optimum LOP, AP provides a 1 dB Q-factor enhancement over FS-DBP. This can be explained as follows: (1) The high peak-to-average power ratio (PAPR) increases the inter-subcarrier XPM among 16-QAM subcarriers that is transformed into time uncertainty through fiber transmission by CD, making the nonlinear interaction appear random. Hence, it is harder for FS-DBP to fully compensate for such a stochastic nonlinear phenomenon. (2) Residual transmitter MZM-based nonlinearity influences the constellation diagram, in which DBP cannot offer any benefit. (3) It is worth noting that parametric noise amplification which affects the system performance in long-haul transmission [2,7], can result in additional stochastic nonlinear distortions. However, there is no evidence that AP clustering can partially tackle this effect. To gain a vibrant representation on the performance benefit of AP regarding point (1), Fig. 4(c) is plotted to show the Q-factor distribution for AP, FL, FS-DBP, and V-NLE against

the middle OFDM subcarriers, which suffer the most from inter-subcarrier nonlinearities. It is evident that AP can further improve the tolerance to inter-subcarrier nonlinearities in the center of the frequency CO-OFDM band compared to other schemes. For instance, for subcarrier#110, AP improves the Q-factor by 2.1 dB over FS-DBP/FL and 3.2 dB in comparison to V-NLE. This confirms the fact that FS-DBP is ineffective in fully compensating inter-subcarrier nonlinearities. On the other hand, AP also outperforms all benchmark algorithms thus improving symbol detection at low powers by simply combating more successfully transceiver imperfections such as residual IQ time-skews and power-imbalances. In WDM-QPSK, AP tackles more effectively inter-channel nonlinearities and has a clear benefit over the entire range of LOPs as shown in Figs. 4(b), (d). At the optimum LOP per channel of -5 dBm, AP increases the Q-factor by 1.4 dB compared to FS-DBP, since the latter algorithm is only valuable for compensation of intra-channel nonlinearities. This is also evident from the illustrated received QPSK constellation diagrams in Fig. 5, for the WDM system at optimum LOP per channel of -5 dBm for cases: (a) w/o NLE, (b) AP clustering, and (c) FS-DBP. The corresponding Q-factors are directly related to Fig. 4(b) at 6.92, 10, and 8.6 dB, respectively. Note that Fig. 5(b) has been plotted with a different colour per cluster to illustrate the nonlinear decision mapping. It is evident that FS-DBP equalizes the signal in terms of intra-channel nonlinearities reduction. However, when AP is applied on the received constellation diagram w/o NLE, its overlapping clustering ability results in better decision de-mapping, thus improving the Q-factor even when compared to an equalized signal from FS-DBP. It is worth noting that for each LOP, the number of iterations for AP to convergence deviates. For high powers, up to ~ 18 iterations are required, while at

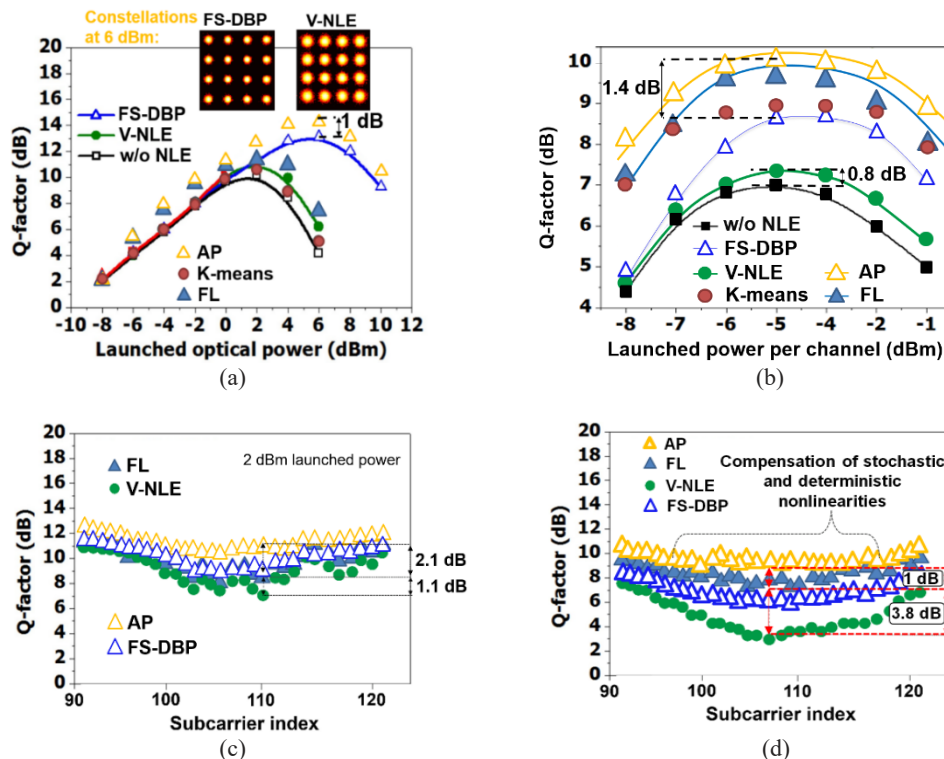


Fig. 4. Q-factor vs. launched optical power (LOP) for CO-OFDM (a) single-channel 16-QAM (2000 km) and (b) WDM-QPSK (3200 km) using different algorithms. The middle subcarriers' performance at optimum LOP is depicted for (c) 16-QAM and (d) WDM-QPSK.

optimum only 10 are needed. This is typically less to state-of-the-art linear adaptive filters used in single-carrier systems requiring >21 taps for long-haul transmission (dual-polarization, DP-QPSK, and 16-QAM standards). In AP, however, the number of operations is $O(i \times k \times n \times 2)$, where i , k , and n are the number of iterations, clusters, and elements, respectively [12]. In terms of complexity, the number of clusters and elements are trivial, while i is hard to predict for different systems and signal modulation format levels due to its interplay with the dispersion structure of the data [12]. A full computational complexity analysis will be provided in future research work.

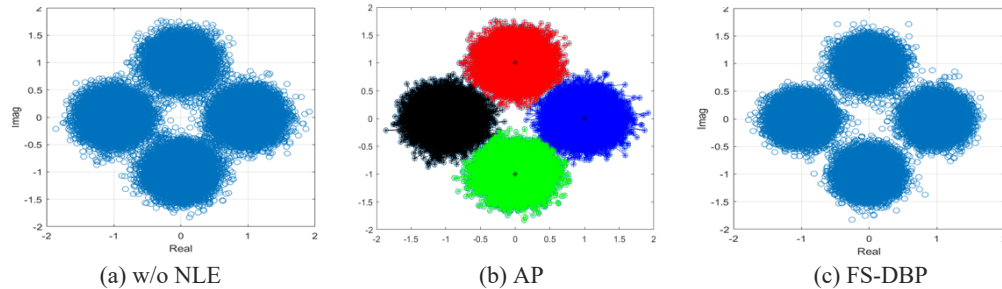


Fig. 5. Received constellation diagrams for WDM-QPSK (3200 km) at optimum LOP per channel of -5 dBm (related to Fig. 4(b)) for: (a) w/o NLE [Q-factor = 6.92 dB], (b) AP clustering [Q-factor = 10 dB], and (c) FS-DBP [Q-factor = 8.6 dB].

5. Simulated comparison between coherent optical OFDM and single-carrier modulation in dual-polarization

In the Section, we compare the performance of the adopted machine learning clustering algorithms between single-carrier and OFDM QPSK/16-QAM modulation in DP. This section essentially addresses the effectiveness of machine learning algorithms on modulated signals with PAPR (i.e., the OFDM case) vs. zero-PAPR signals (i.e., single-carrier case). In particular, we investigate the performance of machine learning for 100 Gb/s QPSK and 200 Gb/s 16-QAM using a DP configuration for both single-carrier and OFDM modulation. For the OFDM case, the transceiver parameters were similar to Section II, but a DP configuration was simulated identical to [13] using two OFDM transmitters for the X- and Y- orthogonal polarizations. The developed OFDM transceivers, as well as the optical transmission, were conducted using VPItransmissionMaker Optical Systems with Matlab co-simulation (electrical domain in Matlab and optical components with SSMF in VPI). The transmitter extinction ratio was kept high (at 35 dB) in order to isolate transmitter imperfections and nonlinearities. To ensure this, we also implemented a pre-distortion of the driving signals for MZM linearization using the well-known $\arcsin(I_s, Q_s) \times V\pi / \pi$, where I_s and Q_s are the in-phase and quadrature components of the complex signal, and $V\pi$ is the peak radio frequency voltage required for π phase change. In the digital CO-OFDM receiver, a zero-forcing 2×2 multiple-input multiple-output (MIMO) equalizer was used that was trained over a sequence of 60 symbols. The DP single-carrier signals were solely simulated in VPItransmissionMaker Optical Systems. The implemented CD compensation algorithm, estimating the reversed linear channel transfer function, was based on the overlap frequency domain equalization [14]. A 2×2 MIMO time-domain equalizer (TDE) was applied in the digital single-carrier coherent receiver using the constant modulus algorithm [15,16] and multi-modulus algorithm [17] for DP-QPSK and DP 16-QAM formats, respectively. The TDE-MIMO comprised a butterfly structure, that enables de-multiplexing of the X and Y polarization tributaries and performed rotation to compensate for the misalignment between the signal and receiver state of polarization.

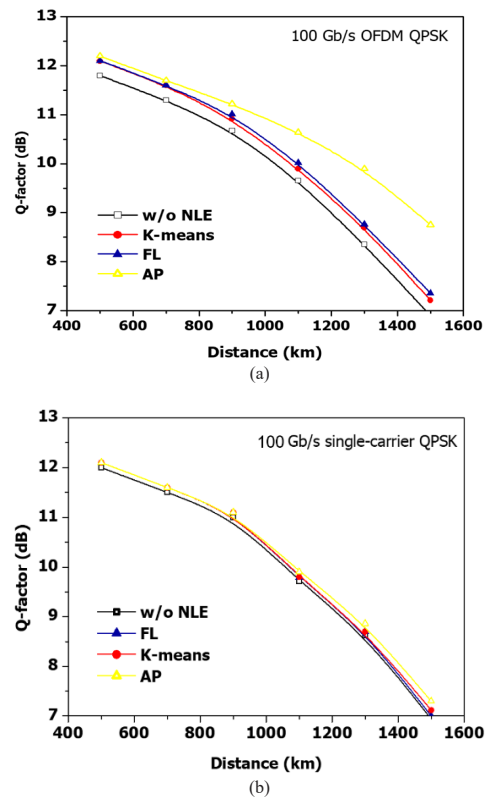


Fig. 6. Q-factor vs. distance for (a) simulated 100 Gb/s CO-OFDM QPSK, and for (b) simulated 100 Gb/s single-carrier QPSK.

Carrier frequency recovery was executed via the 4th power algorithm that was processed in two steps: i) Detection of the maximum peak in the periodogram carried out by an N_s -FFT transform block, with N_s being the number of processed samples when estimating the frequency offset. ii) Compensation by applying a linear phase shift in the time domain. For carrier phase recovery, the sliding window implementation of Viterbi & Viterbi was used when considering DP-QPSK signals [15]. A rectangular windowing filter was applied with the phase error estimated over $2 \times N_{PreSymbols} + 1$ symbols [18]. The $N_{PreSymbols}$ parameter defined the number of pre-symbols used for phase estimation. The phase estimator for 16-QAM signals was based on [17]. Similar to the simulations in Section II, 2^{21} symbols (from an unrepeated pseudo-random binary sequence, PRBS) were used and hard-decision decoding was considered using the Monte-Carlo method. It is worth noting that identical electro-optical noise levels were assumed as in Section II, as well between the single-carrier and OFDM setups in order to provide a fair comparison. Regarding the transmission link simulated in this section, 100 km spans were considered to cover a total distance of up to 2100 km.

In Fig. 6(a), the Q-factor against SSMF transmission distance is plotted for AP, K-means, FL, and LE (i.e., w/o using NLE) for averaged X- and Y-pol. of the 100 Gb/s DP CO-OFDM system, modulated with QPSK across all subcarriers. Results are shown for optimum LOP per distance point. It is evident that over the entire range of transmission distances of interest, all adopted clustering algorithms outperform LE. AP, in particular, shows a significant performance benefit above 800 km of transmission, reaching over 2 dB in Q-factor enhancement observed at 1500 km, while significantly outperforming both K-means and FL. This is because in CO-OFDM we

have higher accumulated inter-subcarrier FWM and XPM at longer distances, leading to higher nonlinear phase noise and more distorted constellation diagrams. AP advanced overlapping clustering ability is much more powerful in this scenario than FL and K-means. In Fig. 6(b), results are shown for averaged X- and Y-pol of the 100 Gb/s single-carrier DP-QPSK system. In contrast to the CO-OFDM system, all clustering algorithms have marginal performance benefits, with the best still being the adopted AP clustering algorithm. Results designate that AP works better under highly unpredicted accumulated deterministic and stochastic inter-subcarrier nonlinear crosstalk effects that are enhanced by the high PAPR. This is due to the fact that AP overlapping clustering leads to incoherent de-mapping, essentially resulting to a form of overestimation. It is also worth noting that the OFDM results in Fig. 6(a) are slightly degraded compared to the single-carrier ones observed in Fig. 6(b). This is because the impact of self-phase modulation on the latter system is always smaller compared to inter-subcarrier nonlinearities.

In Fig. 7, we repeat the simulation process for double signal capacity, using DP-16QAM for both OFDM and single-carrier modulation. Similar to Fig. 6, AP shows a significant Q-factor enhancement only for the OFDM case for distances higher than 1500 km. The highest Q-factor difference compared to LE is observed at 2100 km reaching up to 1.9 dB. This means that compared to Fig. 6 for QPSK when adopting a higher signal modulation format order, the soft-decision boundaries provided by AP are not similarly effective. This also explains the non-smooth performance curves in Fig. 7.

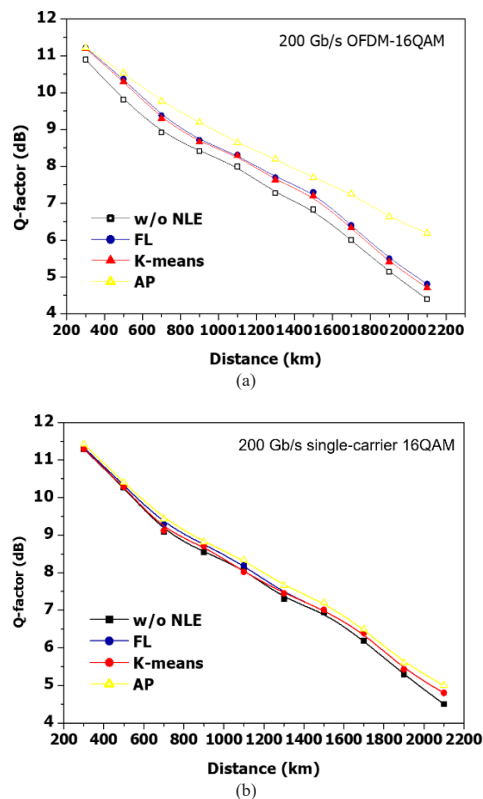


Fig. 7. Q-factor vs. distance for (a) simulated 200 Gb/s CO-OFDM 16-QAM, and for (b) simulated 200 Gb/s single-carrier 16-QAM.

6. Conclusion

AP clustering was experimentally demonstrated in 16-QAM single- and QPSK multi-channel CO-OFDM at 2000 and 3200 km, respectively. We reported significant performance benefits over benchmark deterministic algorithms and machine learning clustering approaches, reaching up to about 4 dB gain in LOP margin. AP effectively tackled nonlinear phase noise induced from both inter-subcarrier/channel FWM/XPM and transmitter nonlinearities. Simulated results indicate transparency of AP for higher-order signal modulation formats and confirmed that it is more effective in a multi-carrier modulation structure. In that sense, AP could benefit state-of-the-art multicarrier schemes such as Nyquist-WDM, while we demonstrated that seamless transferring to DP modulation is feasible.

Funding. Bundesministerium für Bildung und Forschung (16KIS1230); FP7 People: Marie-Curie Actions (623515); Horizon 2020 Framework Programme (713567).

Acknowledgments. We thank Sterlite Techn. (India), and Aston University (UK) for providing experimental OFDM traces. In particular, we acknowledge from Aston Institute of Photonics Technologies (AIPT) the support of Andrew D. Ellis, Nick J. Doran, Son T. Le (now with Nokia Bell-Labs, USA), and Mary E. McCarthy (now with Lumentum, UK).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. E. Temprana, E. Myslivets, BP-P Kuo, L. Liu, V. Ataie, N. Alic, and S. Radic, "Overcoming Kerr-induced capacity limit in optical fiber transmission," *Science* **348**(6242), 1445–1448 (2015).
2. I. D. Phillips, M. Tan, M. F. C. Stephens, M. E. McCarthy, E. Giacomidis, S. Sygletos, P. Rosa, S. Fabbri, S. T. Le, T. Kanesan, S. K. Turitsyn, N. J. Doran, P. Harper, and A. D. Ellis, "Exceeding the nonlinear-Shannon limit using Raman laser based amplification and optical phase conjugation," in *Proc. OFC/NFOEC*, San Francisco, CA, USA, p. M3C.1.
3. G. Gao, J. Zhang, and W. Gu, "Analytical evaluation of practical DBP-based intra-channel nonlinearity compensators," *IEEE Photonics Technol. Lett.* **25**(8), 717–720 (2013).
4. 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," *Nat. Photonics* **7**(7), 560–568 (2013).
5. S. T. Le, M. E. McCarthy, N. M. Suibhne, M. A. Z. Al-Khateeb, E. Giacomidis, N. Doran, A. D. Ellis, and S. K. Turitsyn, "Demonstration of phase-conjugated subcarrier coding for fiber nonlinearity compensation in CO-OFDM transmission," *J. Lightwave Technol.* **33**(11), 2206–2212 (2015).
6. E. Giacomidis, I. Aldaya, M. A. Jarajreh, A. Tsokanos, S. T. Le, F. Farjady, Y. Jaouën, A. D. Ellis, and N. J. Doran, "Volterra-based reconfigurable nonlinear equalizer for dual-polarization multiband coherent OFDM," *IEEE Photonics Technol. Lett.* **26**(14), 1383–1386 (2014).
7. E. Giacomidis, A. Matin, J. Wei, N. J. Doran, L. P. Barry, and X. Wang, "Blind Nonlinearity Equalization by Machine-Learning-Based Clustering for Single- and Multichannel Coherent Optical OFDM," *J. Lightwave Technol.* **36**(3), 721–727 (2018).
8. M. A. Jarajreh, E. Giacomidis, I. Aldaya, S. T. Le, A. Tsokanos, Z. Ghassemlooy, and N. J. Doran, "Artificial neural network nonlinear equalizer for coherent optical OFDM," *IEEE Photonics Technol. Lett.* **27**(4), 387–390 (2015).
9. E. Giacomidis, I. Aldaya, J. L. Wei, C. Sanchez, H. Mrabet, and L. P. Barry, "Affinity propagation clustering for blind nonlinearity compensation in coherent optical OFDM," in *Proc. CLEO*, San Jose, CA, USA, 2018, p. STh1C.5.
10. B. J. Frey and D. Dueck, "Clustering by Passing Messages Between Data Points," *Science* **315**(5814), 972–976 (2007).
11. Y. Ma, M. Peng, and W. Xue, X. Ji, "A dynamic affinity propagation clustering algorithm for cell outage detection in self-healing networks," in *Proc. WCNC*, (IEEE, Shanghai, China, 2013), 2266–2270.
12. E. Giacomidis, Y. Lin, J. Wei, I. Aldaya, A. Tsokanos, and L. P. Barry, "Harnessing machine learning for fiber-induced nonlinearity mitigation in long-haul coherent optical OFDM," *MDPI Future Inter.* **11**(1), 2 (2018).
13. E. Giacomidis, M. A. Jarajreh, S. Sygletos, S. T. Le, F. Farjady, A. Tsokanos, A. Hamié, E. Pincemin, Y. Jaouën, A. D. Ellis, and N. J. Doran, "Dual-polarization multi-band optical OFDM transmission and transceiver limitations for up to 500 Gb/s uncompensated long-haul links," *Opt. Express* **22**(9), 10975–10986 (2014).
14. R. Kudo, T. Kobayashi, K. Ishihara, Y. Takatori, A. Sano, and Y. Miyamoto, "Coherent Optical Single Carrier Transmission Using Overlap Frequency Domain Equalization for Long-Haul Optical Systems," *J. Lightwave Technol.* **27**(16), 3721–3728 (2009).
15. S. J. Savory, "Digital filters for coherent optical receivers," *Opt. Express* **16**(2), 804–817 (2008).

16. E. Pincemin, N. Brochier, M. Selmi, O. Z. Chahabi, P. Ciblat, and Y. Jaouën, "Novel Blind Equalizer for Coherent DP-BPSK Transmission Systems: Theory and Experiment," *IEEE Photonics Technol. Lett.* **25**(18), 1835–1838 (2013).
17. H. Louchet, K. Kuzmin, and A. Richter, "Improved DSP algorithms for coherent 16-QAM transmission," in *Proc. ECOC*, Brussels, Belgium, Sept. 2008, paper Tu.1.E.6.
18. M. Selmi, Y. Jaouen, and P. Ciblat, "Accurate digital frequency offset estimator for coherent PolMux QAM transmission systems," in *Proc. ECOC*, Vienna, Austria, Sept. 2009, paper P3.08.