K-means clustering for mitigation of nonlinear phase noise in digital coherent optical systems using 16-QAM modulation format

Lúcio N. Borges, Camila B. Costa, Rômulo de Paula, Marcelo L. F. Abbade, and Ivan Aldaya

***Abstract*— This work analyzes the use of the K-means clustering algorithm to mitigate nonlinear phase noise in single-span coherent systems, such as long-reach passive optical networks (LR-PONs). Simulations revealed that for a 100-km LR-PON employing 16-ary quadrature amplitude modulation (QAM) and considering a 1:64 splitting ratio, the adoption of K-means with K-means++ initialization achieves an optimum bit error ratio (BER) of $6.3 \cdot 10^{-4}$, whereas employing maximum likelihood, 10^{-3} is obtained. We also show that in order to achieve this performance improvement in 90% of the cases, K-means requires only 2,000 symbols.**

***Keywords*— Passive optical networks; Nonlinear compensation; Nonlinear phase noise; K-means.**

I. Introduction

Since the 1980s, different research groups have proposed solutions to improve the performance of optical communication systems [1]. With this objective, coherent optical detection was explored, but the development of erbium doped fiber amplifiers (EDFAs) made research to remain restricted to systems based on intensity modulation with direct detection (IM/DD). In particular, the combination of lasers' phase noise and fluctuation of the state of polarization (SoP) of the received signal prevented the adoption of coherent detection [2]. However, in mid 2000s, coherent systems regained attention thanks to new advances in digital signal processors (DSPs). Thus, the development of digital coherent systems enabled the implementation of new compensation techniques and adoption of more efficient modulation formats, such as M-ary phase shift keying (m-PSK) and quadrature amplitude modulation (QAM). The enhanced spectral efficiency, then, led to a significant increase in data transmission capacity and optimized the network efficiency [3] [4].

Clearly, the migration to coherent systems represented a major breakthrough in data transmission rates. Digital coherent systems, in addition to present an improved receiver sensitivity, they enable to recover both phase and polarization information, thus allowing compensation of chromatic dispersion (CD) [5], polarization mode dispersion (PMD) [6], rotation of the SoP [7] , and linear phase noise [8] in a novel and more efficient way. Nonlinear distortion, nevertheless, remains as an unresolved issue and, its interplay with additive noise still poses an upper bound to the capacity-range product of the network. In coherent optical communications systems, Kerr effect is commonly claimed as the main source of nonlinear distortion [9]. It is well known that this effect further causes self-phase modulation (SPM), cross-phase modulation (XPM), and four-wave mixing (FWM), but in single-channel links, such as broadcasted long-reach passive optical networks (LR-PONs), only SPM is present [10]. The main effect of SPM is a variation of the refractive index according to the power of the transmitted signal. Consequently, in modulation formats with multiple amplitude levels, such as 16-QAM, the points on the periphery of the constellation suffer a greater rotation, leading to a characteristic spiral-like constellation [4]. In addition, even if CD is compensated at the receiver side, some intersymbol interference (ISI) occurs along the fiber. This ISI causes adjacent symbols to interact nonlinearly via SPM, as the rotation depends on the instantaneous intensity. As a result, the nonlinear distortion of the constellation seems stochastic and often treated as a nonlinear noise. Indeed, it is usually referred as nonlinear phase noise (NPN) or Gordon-Mollenauer effect [11].

To mitigate nonlinear distortion, including nonlinear phase noise, several techniques have been proposed, both in the optical and electrical domains. Among the explored optical techniques, we can mention twin-waves [12] and mid-span conjugation [13]. Unfortunately, the optical methods are complex and difficult to implement in the field. On the other hand, electrical nonlinear compensation avails the flexibility that electronics offers. The first implementations of this type of compensators were based on model inversion methods. For instance, digital backward propagation (DBP) [9] and inverse Volterra-series transfer function (IVSTF) [14] [15]. Inversion methods, however, require a huge amount of operations for impairment correction, which hinders its adoption in real-time systems. In order to reduce the computational complexity, machine learning algorithms have been recently proposed. In particular, support vector machines (SVM) [16], artificial neural networks (ANNs) [17], and K-nearest-neighbors (KNN) [18] have already been reported. All these approaches are supervised classification techniques, meaning that they require a training sequence in order to find the model parameters that best match the system to be equalized. Alternatively, some unsupervised classification algorithms, also known as clustering, have been studied. For example, in [19] and [20] expectation maximization and histogram-based-clustering are employed, respectively. Due to its low complexity in the test stage, K-means is another clustering algorithm that has been

Campus São João da Boa Vista, Universidade Estadual Paulista "Júlio de Mesquita Filho" - Unesp, São João da Boa Vista - SP, e-mails: lucio.borges@unesp.br, cb.costa@unesp.br, romulo.junior314@yahoo.com.br, marcelo.abbade@unesp.br, and ivan.aldaya@unesp.br.

studied to compensate nonlinear distortion [21] [22]. However, in [21], it is not clear whether the algorithm is compensating the fiber nonlinearities or other residual linear impairments. On the other hand, in [22], the algorithm is not tested for different launch optical power levels, which makes difficult to assess the improvement of the technique at optimum configuration. Furthermore, in [21], K-means is initialized either using a training sequence or setting positions of the centroids according to a rectangular grid, making automation far from trivial. On the other hand, [22] relays on random initialization that, as we will see in the next Section, may result in suboptimal clustering and requires multiple tentatives. This issue is particularly critical as the number of cluster k increases. Among the different initialization methods, K-means++ has been proposed as a solution.

In this paper, we evaluate the performance of K-means with K-means++ initialization in optical systems affected by the combined effect of NPN and receiver additive noise, which to our best knowledge has not yet been accomplished. We present simulation results of coherent 16-QAM single-span links where the effect of NPN is mitigated using this approach. Numerical results reveal that considering a block with 2,000 symbols or more, the bit error ratio (BER) results are approximately constant.

II. K-MEANS CLUSTERING CLASSIFICATION

As mentioned, K-means clustering is an efficient and simple unsupervised classification algorithm. Therefore, it does not require data to be manually labeled and it dispenses the reward method needed in reinforcement learning. The algorithm creates *K* similar entries grouped automatically by just comparing the characteristics of received symbols. In K-means, each cluster is associated to a centroid, which verify that the distance between all the symbols to be classified to their associated centroids is minimal. In order to get the position of the centroids, the Lloyds algorithm is usually adopted, which consists of the following steps [23]:

- **Step 1:** Initialization of the centroids to initial positions μ_j;
- **Step 2:** Classify the points by associating them to the closest centroid. This is done by finding the centroid that minimizes the distance to the given symbol:

$$d_i = min_j\{d(x_i, \mu_j)\} \ \forall i \text{ and } 1 \leq j \leq K, \qquad (1)$$

  where $d(\cdot)$ represents a distance operator (typically the Euclidean distance). The points associated to a given centroid constitute a cluster;
- **Step 3:** Find the new positions of the centroids by calculating the center of the data associated to each cluster according to:

$$\mu'_j = \frac{1}{N_j}\sum_{i=1}^{N_j} x_{i,j} \text{ for } 1 \leq j \leq K, \qquad (2)$$

  N_j and $x_{i,j}$ being the number of elements associated to cluster j and the elements of this cluster, respectively;
- **Step 4:** If the calculated centroids μ'_j are different from the previous centroids μ_j, update the centroids by equaling $\mu_j = \mu'_j$ and return to Step 2.

At the end of this process, the algorithm outputs the k centroids as well as the points associated to each one. It is worth noting that the complexities of Step 2 and Step 3 increase as the number of data to be classified and the number of clusters do. Hence, the complexity of the overall process also depends on these parameters. Even if some heuristic algorithms have demonstrated such fast convergence and a similar computational cost [23], we will implement the aforementioned Lloyds algorithm, as this is the most popular and simpler approach.

A well-known disadvantage of K-means is its sensitivity to initial selection of cluster positions. Thus, random initialization may converge to a local minimum, specially for a large number of clusters [24]. One of the solutions to this issue is the repetition of the process with different random initial solutions, but this approach increases significantly the computational cost. Alternatively, smarter initialization methods has been proposed, for instance: Simple Cluster Seeking method [25]; Global K-means method [26]; Cluster center initialization algorithm [27]; Hierarchical K-means [28]; and K-means++ [24]. K-means++ is a heuristic initialization algorithm that chooses k data points separated by large distances. In order to find these points the following process is performed:

- **Step 1:** Choose μ_1 randomly from the set of data points to be classified x_i;
- **Step 2:** Compute the distance d_i from all the data points x_i to the already chosen centroid(s);
- **Step 3:** Choose the next centroid μ_j by selecting the data point with the maximum probability, which is given by:

$$\text{Prob} = \frac{[d_i]^2}{\sum_{\forall i}[d_i(x)]^2}; \qquad (3)$$

- **Step 4:** Repeat Step 2 and 3 until k centroids are found.

In order to see the importance of proper initialization, Fig. 1(a) shows an unsuccessful classification caused by an incorrect random initialization, whereas Fig. 1(b) represents a classification employing K-means++ initialization. In addition

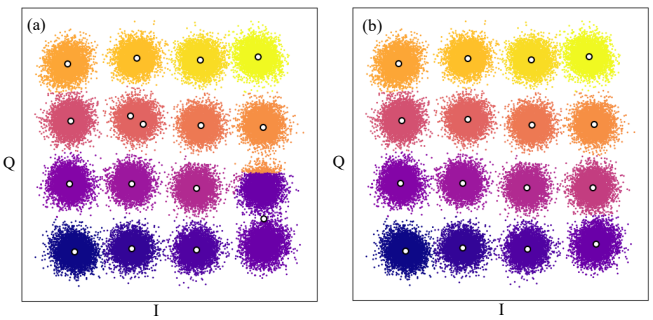

Fig. 1. Constellation diagram for comparison between K-means with random initialization and K-means++ initialization. (a) Diagram with random startup and centroids for each cluster. (b) Diagram with K-means++ initialization, with the same data. For both cases, the found centroids are superposed.

to better classification performance, a smarter initialization scheme typically results in lower iterations to converge, thus requiring less operation count. In the case showed in Fig. 1, for example, the number of iterations for K-means was 19 and the iterations using K-means with K-means++ initialization was 8.

III. Simulation Setup

In this work, we used Matlab to perform the modulation, demodulation, and linear impairment compensation. VPI Transmission Maker software was employed to simulate the electrical-to-optical and optical-to-electrical conversions, as well as the propagation through the fiber. The nonlinear impairment compensation based on K-means clustering was implemented in Python. To test the capabilities of K-means with K-means++ initialization in a system strongly affected by NPN, we simulated a coherent LR-PON operating at single polarization and employing 16-QAM modulation format. The signal quality was assessed through the BER, which was found by counting the number of errors. In addition, the BER using maximum likelihood (ML) detection was calculated, as this detection method is commonly used as benchmark.

The employed simulation setup is shown in Fig. 2. In the transmitter, a continuous wave (CW) laser diode (LD) with a 100-kHz linewidth and a power of 1 mW operating at 1550 nm was externally modulated employing a dual parallel Mach–Zehnder modulator (DP-MZM). The modulator was driven by the in-phase and quadrature components of a 56-Gbps 16-QAM signal with a bandwidth of 10.5 GHz, which was limited by a fourth-order Bessel filter. The modulated signal was then amplified by an erbium-doped fiber amplifier (EDFA), whose output power was set to 10 mW. It is important to note that since the input power to the EDFA was relatively high, the aggregated amplified spontaneous emission (ASE) noise was small. In addition, we used a variable optical attenuator (VOA) to sweep the power launched into the fiber from 2 to 10 mW. The distribution network was composed of a first link of 80 km standard single mode fiber (SSMF), an 18-dB attenuator representing a 1×64 splitter, and a second link of 20-km SSMF. Both SSMFs were simulated using split step Fourier method (SSFM) with adaptive steps. At the receiver front-end, the polarization of the incoming signal was first controlled by a dynamic polarization tracker (DPT) to minimize the polarization mismatch with the local oscillator laser. Then, the signal was mixed in a 90^{o} optical hybrid network with a 1-mW CW laser. The outputs of the hybrid network were photodetected, filtered, and differentially amplified to obtain the in-phase and quadrature components of the received signal. Table I summarizes the main simulation parameters.

Once in the electrical domain, the in-phase and quadrature signals were downsampled to get 4 samples per symbol, which emulates the analog-to-digital conversion (ADC). The CD compensation was performed in the frequency domain. Subsequently, the time synchronization was carried out using the cross-correlation maximization method employing a pilot alternating sequence of 64 symbols. This pilot sequence was also used to correct the initial phase rotation. The signal was further downsampled to a single sample per symbol, after which the rotation caused by lasers' phase noises was corrected by blind phase search with blocks of 32 symbols [29]. It is important to note that for spans of 100 km, the accumulated PMD is much smaller than the symbol period and, consequently, we did not have to implement PMD compensation. Finally, the obtained symbols were classified using both ML and K-means clustering with K-means++ initialization.

TABLE I
Simulation parameters.

Parameters		
Signal	Modulation format	16-QAM
	Bit Rate	56 Gbps
	Number of synchronization symbols	64
	Order of eletrical TX filter (Bessel)	4^{th}
System	Laser power	1 mW
	Laser linewidth	0.5 MHz
	MZM loss	6 dB
	Amplifier noise figure	4 dB
	Amplifier gain	20 dB
	Attenuator	20 dB
	PD responsivity	1 W/A
	PD thermal noise density	10 pA/$\sqrt{\text{Hz}}$
	Fiber PMD	3.16 fs/$\sqrt{\text{km}}$
	Fiber attenuation	0.2 dB
	Fiber CD	16 ps/nm/km
	Fiber lenghts	80 km and 20 km
	Fiber effective area	80 μm^2
	Nonlinear coefficient	1.3 $W^{-1}{\cdot}km^{-1}$
	Eletrical RX filter order	4
	Eletrical filter bandwidth	10.5 GHz
Simulation	Number of simulated symbols	65000
	Number of simulated bits	260000
	Sampling rate	8.96×10^{11} s^{-1}

IV. Results and Discussion

In order to evaluate the performance of the proposed algorithm, in Fig. 3(a), we show the BER curves for ML and K-means classifications in terms of the launch optical power. We can observe that, at low power levels, when the system is limited by the additive noise of the photodetectors, the performances of both techniques are very similar. When we compare the BER obtained by ML and K-means classification for power levels greater than 3 mW, the latter presents better performance. Therefore, we can conclude that K-means is indeed compensating nonlinear distortion and not any other residual linear impairments. In particular, the optimum BER using ML is 10, obtained at a launch optical power of 6 mW, whereas when K-means with K-means++ initialization is employed, an optimum BER of $6.3\cdot10^{-4}$ is achieved at the same optimal power. In addition, if we classify the received symbols using K-means, we can increase the launch optical power up to 8 mW, maintaining the optimum BER of ML detection.

To understand the operation of K-means in compensating nonlinear distortion, specially SPM, we analyze the constellations after they were classified by ML and K-means methods.

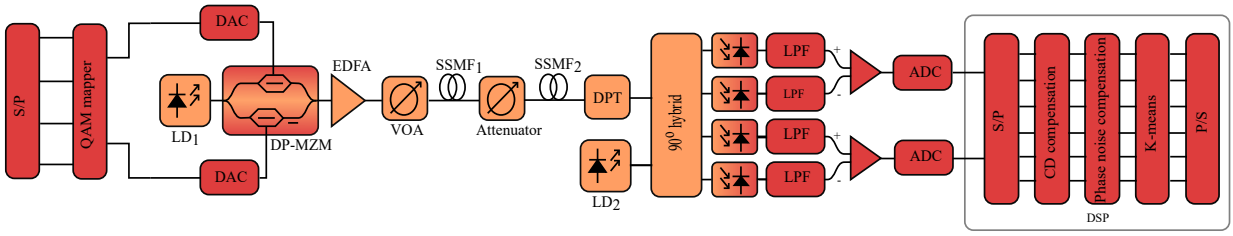

Fig. 2. Block diagram of the simulated coherent LR-PON system. S/P: serial-to-parallel converter; DAC: digital-to-analog converter; LD: laser diode; DP-MZM: dual parallel Mach–Zehnder modulator; EDFA: erbium-doped fiber amplifier; VOA: variable optical attenuator; SSMF: standard single mode fiber; DPT: Dynamic polarization tracker; LPF: low-pass filter; ADC: analog-to-digital converter; and P/S: parallel-to-serial converter.

For the sake of illustration, we only present the obtained constellation for 2, 6, and 10 mW for ML, Fig. 3(b-d), and for the same power levels using K-means, Fig. 3(e-g). When comparing the constellations at low power levels presented in Fig. 3(b) and (e), both detection methods results in similar rectangular decision regions, which explains the same BER performance. However, when the power increases, the SPM distorts the constellation, and non-rectangular decision regions lead to an improved BER. For instance, at medium power values, as shown in Fig. 3(c) and (f), the centroids found by K-means were not arranged in a rectangular geometry. The resulted irregular decision regions better match the dispersion of the points, resulting in a better classification of the data and improving the BER compared to the use of ML. This behavior is even more apparent at higher power levels, as we can see in Fig. 3(d) and (g).

Due to the required heavy signal processing, modern optical communication systems may present significant latency and power consumption. Therefore, it is necessary to reduce the operations required by each processing block. In the case of nonlinearity compensation based on clustering, the amount of operations and, consequently, the latency and power consumption, depend on the size of the data block to be classified. Thus, excessively large blocks will incur in prohibitively high latency and computational load, whereas processing short blocks will degrade the performance of the classification process. Therefore, it is important to find a trade-off between computation complexity and performance. In order to properly choose the block size, a second analysis was performed in this work, testing the efficiency of the algorithm for different data block sizes. To perform this study, we randomly selected data blocks with different sizes. As this set is random, so it is the classification process and several classification trials must be carried out to estimate the ensemble behavior of the clustering process. Thus, for each block size we randomly collected 100 samples from the complete set of symbols. K-means++ was applied to each of these samples and the results of these classifications are presented in Fig. 4. The scattered dots in salmon color represent the computed BER for each block sample when data block sizes were swept from 100 to 25,000 data. The brown curve shows the average of the individual BERs, whereas the gray area indicates the region with the 90% lowest BERs. As can be observed, for block sizes smaller than 2,000 symbols, the variation of the performance is significant, which indicates that poor classification may occur

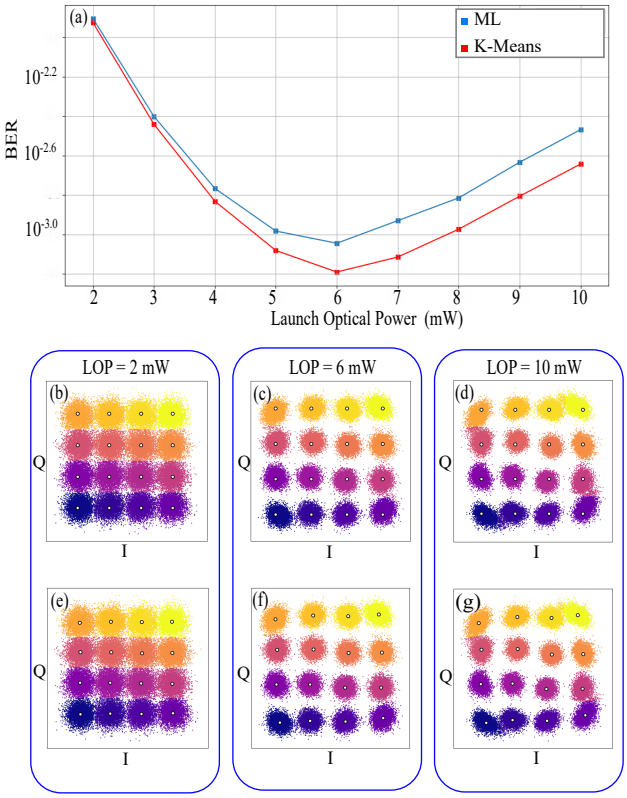

Fig. 3. Performance of ML and K-means. (a) BER curves in terms of launch optical power for ML and K-means with K-means++ initialization. (b-d) Constellation diagrams for 2, 6, and 10 mW classified using ML and (e-g) Constellation diagrams for 2, 6, and 10 mW when K-means is employed for classification.

frequently. As the block size increases, the variance of BER values reduces and the performance of the classification is more consistent. In particular, for set sizes above 2,000 data, we have 90% of the computed BER below $6.3 \cdot 10^{-4}$. In the same figure, we can see that for sets larger than 15,000 data, BER values will be less than $5.01 \cdot 10^{-4}$ in the 90% of cases.

V. Conclusions

In this paper we analyzed the performance and block size requirements of K-means clustering for the compensation of NPN in coherent LR-PONs employing 16-QAM. Simulation results reveal that the optimum BER can be decreased

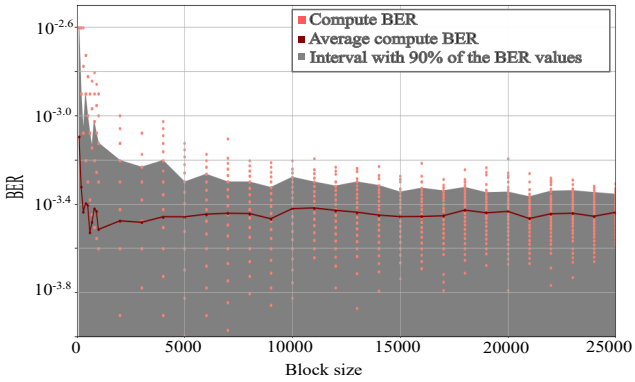

Fig. 4. BER in terms of the processed block size. The points correspond to BER calculated for different block samples, the solid line represents the ensemble average of the sample BERs, and the gray area indicates the region where the lowest 90% of the calculated BERs are contained.

from 10^{-3} when using ML to $6.3 \cdot 10^{-4}$ by employing K-means clustering with K-means++ initialization. In addition, we showed that if K-means is adopted, the launch optical power can be increased from 6 mW to 8 mW while keeping the optimum performance achievable by ML. In order to attain this performance improvement, K-means requires to operate over blocks of at least 2,000 symbols. This block size results in a block latency of 0.14 μs for a 56-Gbps 16-QAM system. These results then show the potential of K-means with K-means++ initialization for compensation of NPN in real-time implementations. It is expected this potential to be even more significant in higher order modulation formats, such as 64-QAM, which is left as future work.

ACKNOWLEDGMENTS

The authors thank the National Council for Scientific and Technological Development (CNPq, grant numbers 432303/2018-9 and 311035/2018-3) and the São Paulo Research Foundation (FAPESP, grant 2018/25339-4).

REFERENCES

[1] G. P. Agrawal, "Optical communication: its history and recent progress," in *Optics in Our Time*. Springer, Cham, 2016, pp. 177–199.

[2] K. Kikuchi, "Coherent optical communication systems," in *Optical Fiber Telecommunications VB*. Elsevier, 2008, pp. 95–129.

[3] S. Tsukamoto, D.-S. Ly-Gagnon, K. Katoh, and K. Kikuchi, "Coherent demodulation of 40-Gbit/s polarization-multiplexed QPSK signals with 16-GHz spacing after 200-km transmission," in *Optical Fiber Communication Conference*. Optical Society of America, 2005, p. PDP29.

[4] K. Kikuchi, "Fundamentals of coherent optical fiber communications," *Journal of Lightwave Technology*, vol. 34, no. 1, pp. 157–179, 2015.

[5] A. Green, P. Mitra, and L. Wegener, "Effect of chromatic dispersion on nonlinear phase noise," *Optics Letters*, vol. 28, no. 24, pp. 2455–2457, 2003.

[6] H. Sunnerud, C. Xie, M. Karlsson, R. Samuelsson, and P. A. Andrekson, "A comparison between different PMD compensation techniques," *Journal of Lightwave Technology*, vol. 20, no. 3, p. 368, 2002.

[7] N. Ekanayake and H.Herath, "Effect of nonlinear phase noise on the performance of M-ary PSK signals in optical fiber links," *Journal Lightware Technology*, vol. 31, no. 3, pp. 447–454, 2013.

[8] K.-P. Ho and J. M. Kahn, "Electronic compensation technique to mitigate nonlinear phase noise," *Journal of Lightwave Technology*, vol. 22, no. 3, p. 779, 2004.

[9] C.-Y. Lin, R. Asif, M. Holtmannspoetter, and B. Schmauss, "Nonlinear mitigation using carrier phase estimation and digital backward propagation in coherent QAM transmission," *Optics Express*, vol. 20, no. 26, pp. B405–B412, 2012.

[10] A. Demir, "Nonlinear phase noise in optical-fiber-communication systems," *Journal of Lightwave Technology*, vol. 25, no. 8, pp. 2002–2032, 2007.

[11] J. P. Gordon and L. F. Mollenauer, "Phase noise in photonic communications systems using linear amplifiers," *Optics Letters*, vol. 15, no. 23, pp. 1351–1353, 1990.

[12] S. Le, M. McCarthy, S. Turitsyn, I. Phillips, D. Lavery, T. Xu, P. Bayvel, and A. Ellis, "Optical and digital phase conjugation techniques for fiber nonlinearity compensation," in *Opto-Electronics and Communications Conference (OECC)*. IEEE, 2015, pp. 1–3.

[13] K.-P. Ho, "Mid-span compensation of nonlinear phase noise," *Optics Communications*, vol. 245, no. 1-6, pp. 391–398, 2005.

[14] E. Giacoumidis, I. Aldaya, V. Vgenopoulou, N. Doran, and Y. Jaouën, "Impact of reduced complexity inverse volterra series transfer function-based nonlinear equalizer in coherent OFDM systems for next-generation core networks," in *2013 15th International Conference on Transparent Optical Networks (ICTON)*. IEEE, 2013, pp. 1–4.

[15] Z. Pan, B. Châtelain, M. Chagnon, and D. V. Plant, "Volterra filtering for nonlinearity impairment mitigation in DP-16QAM and DP-QPSK fiber optic communication systems," in *National Fiber Optic Engineers Conference*. Optical Society of America, 2011, p. JThA040.

[16] D. Wang, M. Zhang, Z. Li, Y. Cui, J. Liu, Y. Yang, and H. Wang, "Nonlinear decision boundary created by a machine learning-based classifier to mitigate nonlinear phase noise," in *European Conference on Optical Communication (ECOC)*, Sep. 2015, pp. 3–16.

[17] M. A. Jarajreh, E. Giacoumidis, I. Aldaya, S. T. Le, A. Tsokanos, Z. Ghassemlooy, and N. J. Doran, "Artificial neural network nonlinear equalizer for coherent optical OFDM," *IEEE Photonics Technology Letters*, vol. 27, no. 4, pp. 387–390, 2015.

[18] D. Wang, M. Zhang, M. Fu, Z. Cai, Z. Li, H. Han, Y. Cui, and B. Luo, "Nonlinearity mitigation using a machine learning detector based on k-Nearest Neighbors," *IEEE Photonics Technology Letters*, vol. 28, no. 19, pp. 2102–2105, 2016.

[19] D. Zibar, O. Winther, N. Franceschi, R. Borkowski, A. Caballero, V. Arlunno, M. N. Schmidt, N. G. Gonzales, B. Mao, and Y. Ye, "Nonlinear impairment compensation using expectation maximization for dispersion managed and unmanaged PDM 16-QAM transmission," *Optics Express*, vol. 20, no. 26, pp. B181–B196, 2012.

[20] I. Aldaya, E. Giacoumidis, G. de Oliveira, J. Wei, J. L. Pita, J. D. Marconi, E. A. M. Fagotto, L. Barry, and M. L. F. Abbade, "Histogram based clustering for nonlinear compensation in long reach coherent passive optical networks," *Applied Sciences*, vol. 10, no. 1, p. 152, 2020.

[21] M. G. Junfeng Zhang, Wei Chen and G. Shen, "K-means-clustering-based fiber nonlinearity equalization techniques for 64-QAM coherent optical communication system," *Optics Express*, vol. 25, no. 22, pp. 27 570–27 580, 2017.

[22] E. A. Fernandez, J. J. G. Torres, A. M. C. Soto, and N. G. González, "Demodulation of m-ary non-symmetrical constellations using clustering techniques in optical communication systems," in *2016 IEEE Latin American Conference on Computational Intelligence (LA-CCI)*. IEEE, 2016, pp. 1–6.

[23] S. Marsland, *Machine Learning: an Algorithmic Perspective*. CRC Press, 2015.

[24] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," Stanford, Tech. Rep., 2006.

[25] J. T. Tou, "DYNOC-A dynamic optimal cluster-seeking technique," *International Journal of Computer & Information Sciences*, vol. 8, no. 6, pp. 541–547, 1979.

[26] A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," *Pattern Recognition Letters*, vol. 36, no. 2, pp. 451–461, 2003.

[27] S. S. Khan and A. Ahmad, "Cluster center initialization algorithm for k-means clustering," *Pattern Recognition Letters*, vol. 25, no. 11, pp. 1293–1302, 2004.

[28] K. Arai and A. R. Barakbah, "Hierarchical k-means: an algorithm for centroids initialization for k-means," *Reports of the Faculty of Science and Engineering*, vol. 36, no. 1, pp. 25–31, 2007.

[29] T. Pfau, S. Hoffmann, and R. Noé, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for M-QAM constellations," *Journal of Lightwave Technology*, vol. 27, no. 8, pp. 989–999, 2009.