Intelligent Analog Radio Over Fiber aided C-RAN for Mitigating Nonlinearity and Improving Robustness

Yichuan Li
Department of Electronic and Information Engineering
Harbin Institute of Technology (Shenzhen)
Shenzhen, China
liyichuan@hit.edu.cn

Mohammed El-Hajjar
School of Electronics and Computer Science
University of Southampton
Southampton, United Kingdom
meh@ecs.soton.ac.uk

***Abstract*—As a low-cost solution for the 5G communication system, centralised radio access network (C-RAN) has been implemented in the ultra-dense environment, where radio over fiber (RoF) technology can enable reduced operational cost as well as coordinated multi-point (CoMP) despite its less-robustness and reduced system performance. On the other hand, machine learning has been recognised as an efficient method for accelerating the fiber-optic communications with the aid of the advancements of the learning algorithms as well as the available high processing capabilities. In this paper, we propose a supervised learning-aided A-RoF system, where the logistic regression classification is invoked for removing the A-RoF module's need for re-customization and for boosting its performance. As a result, we can adaptively select the modulation format according to the optical power and the RF voltage, where we obtain an enhanced spectral efficiency and dynamic range (DR) by a factor of 4/3 and 19/13, respectively, while the learning network can be updated online.**

***Index Terms*—Analogue radio over fiber (A-RoF), Centralised radio access network (C-RAN), supervised learning, logistic regression classification, fronthaul.**

I. Introduction

At the time of writing, the fiber-based radio access network (RAN) has been evolving for decades for the sake of supporting various services [1]–[3], such as 5G, Wi-Fi, voice or video call, etc., while supporting seamless connections, ultra-low latency and massive data access [1], [4]. The latest RAN architecture defined by the enhanced common public radio interface (eCPRI) employs the digitised radio over fiber (D-RoF) and functional split to reduce the data rate required in the RAN fronthaul by relocating some physical layer function to the remote radio head (RRH) [5]. However, it is at the cost of increased RRH's complexity, cost and less efficient centralised RAN (C-RAN) virtualization [1], [5].

The financial support of Shenzhen Scientific Research Foundation for the introduction of High-Caliber Personnel under Grant JB11409017, the Shenzhen Postdoctoral Science Foundation under Grant JB25501024, the Research Foundation of HITsz under Grant JB45001037 is gratefully acknowledged.

Yichuan Li is with Harbin Institute of Technology (Shenzhen), Shenzhen, China (e-mail: liyichuan@hit.edu.cn)

M. El-Hajjar is with the School of ECS, University of Southampton, SO17 1BJ, United Kingdom (meh@ecs.soton.ac.uk).

In this context, implementing the analogue radio over fiber (A-RoF) technique in the fronthaul system provides a low-cost and bandwidth-friendly solution, since it can eliminate the need for the digital signal processing [1], [4], [5], such as digital-to-analogue/analogue-to-digital converter (ADC/DAC). On the other hand, the A-RoF can reduce the optical bandwidth from 295 GHz needed in the D-RoF to 10 GHz using the A-RoF if a 16×16 multiple-input-multiple-output (MIMO) system is considered [1]. Despite its huge bandwidth- and power- efficiency, the A-RoF based fronthaul remains renounced by the mobile operators due to the following challenges. Firstly, A-RoF are less robust than D-RoF and require re-customisation, which increases the expenditure cost for robustness management and installation fee, which is a priority for the mobile operators [6]. Then, A-RoF experiences reduced performance in terms of error vector magnitude (EVM) compared to the D-RoF, which affects the quality of service of the fronthaul system [1].

Specifically, the joint effects of noise, fiber-link nonlinearity and dispersion challenge the A-RoF's feasibility for the next-generation wireless RAN [7]. Hence, there has been several research projects addressing these challenges [8]. For example, novel optical devices and novel architecture have been investigated for optical noise-reduction by low biasing the optical modulator [8], while the fiber's and optical modulation's non-linearity have been analysed and investigated using multiple fiber transmission in [9] and using linearized optical modulator in [10]. Additionally, overcoming the chromatic dispersion exploiting fiber grating was designed in [11]. On the other hand, advanced optical signal processing including pre-distortion and post-compensation technique [12] and optical phase conjugate [13], [14] have been proposed to increase the linearization of the optical modulator and reduce the fiber's chromatic dispersion. These techniques are designed for lower-order modulations, which is lower than 64 QAM [1], [4] and mainly considering discrete fiber-link impairments, while ignoring the interaction amongest fiber-link nonlinearity and noise as well as the fiber dispersions. Hence these techniques fail to provide a robust fronthaul system, where parameters

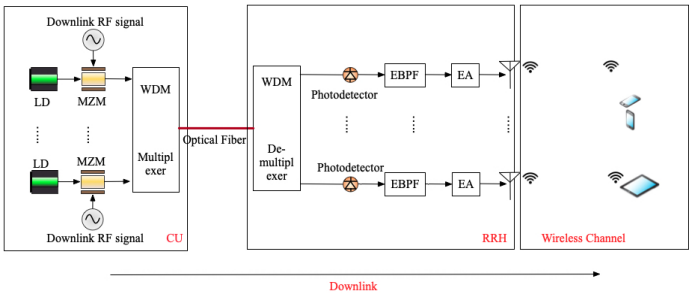

Fig. 1. Conventional A-RoF Fronthaul Downlink. LD: Laser Diode, MZM: Mach-zenhder Modulator, WDM: Wavelength Division Multiplexing, EBPF: Electronic Band-pass Filter.

such as modulation format, the range of optical power and fiber link, and the amplitude of radio frequencies should be adapted for various contexts.

In recent years, machine learning has been actively applied in areas of optical communications such as resource allocation, signal demultiplexing and decoding, linear and nonlinear equalizing [15], [16]. Furthermore, a comprehensive review of machine learning for A-RoF is demonstrated in [17], justifying the feasibility of machine learning for improving the A-RoF's quality of service (QoS).

Moreover, due to the implementation of higher-order modulation and wider bandwidth for pursuing higher data rate, the A-RoF fronthaul still suffers from severe fiber-link impairment. *In order to mitigate the fiber-link adversaries, in this paper, in the context of an A-RoF fronthaul downlink using 64 QAM and 256 QAM, and a carrier aggregation of 16 carrier components, we design a supervised learning aided A-RoF C-RAN system, where the central unit (CU) performs the learning algorithm using logistic regression to improve both the capacity and the EVM performance by predicting the optimal modulation format based on the given optical and RF power. Explicitly, we exploit logistic regression classification algorithm for the sake of improving the system robustness and to increase the achievable system capacity of the potentially deployed next-generation A-RoF C-RAN system [1].*

Our contributions are summarised as follows:

1) Instead of using the traditional closed-form equation to solve the classification problem of modulation formats, which suffers from a high computing complexity [18], [19], we propose a logistic regression classification model to predict the EVM performance of the A-RoF fronthaul system based on the RF power and optical power (after the optical source). As a result, the supervised logistic regression learning algorithm can mitigate the impact of the RF power and optical power on the E/O's nonlinearity and fiber nonlinearity.
2) We conceive an A-RoF system, which is capable of increasing the capacity by invoking an learning aided adaptive modulation selection scheme by selecting the higher-order modulation format, while achieving a reliable system performance.

The rest of this paper is organised as follows. We describe the conventional system model in Section II, where we present the conventional A-RoF system architecture, after which we employ the logistic regression classification model for the A-RoF fronthaul downlink in Section III, while the results and analysis are presented in Section IV. Finally, we conclude in Section V.

II. The Conventional A-RoF System Architecture

TABLE I
Simulation Parameter

Parameter	Value
Subcarriers (IEEE 802.11 a)	(Data 48; Pilot 4; Null 1)
Carrier Aggregation	16 Component Carriers (5G NR)
Bandwidth	800 MHz/ per wavelength (antenna)
Subcarrier Spacing	312.5 KHz
Modulation	256 QAM/64QAM
RF Carrier	3 GHz
WDM Spacing	25 GHz (Dense WDM)
Fiber Length	20 Km
Fiber type	Standard Single-mode Fiber
Fiber length	10 km
EVM Requirement (3GPP Rel. 15)	3.5% for 256 QAM; 8% for 64 QAM
Simulation Environment	MATLAB

In this section, we present an overview of the conventional A-RoF systems and its corresponding challenges regarding the dynamic range and EVM performance for various modulation formats.

As mentioned in Section I, our design is based on the A-RoF aided C-RAN architecture, where the connection between the CU and RRH, namely fronthaul, transmits the modulated radio frequency signal via the optical fiber. Fig. 1 presents the architecture for low-cost and high-performance fronthaul solutions [1], [4], where the analogue fronthaul eliminates the requirement for high-speed power-consuming ADC/DAC in the RRH, while the wavelength-division-multiplexing (WDM) aided MIMO signal processing can improve the signal-to-noise ratio by using analogue beamforming, and can also increase the channel capacity by exploiting the spatial multiplexing or combat the multi-path fading by implementing diversity schemes [20], [21]. In Fig. 1, we demonstrate a WDM based A-RoF fronthaul downlink. The A-RoF fronthaul downlink is framed by a central unit (CU) and several RRHs, which is termed as C-RAN, that is a low-cost RAN solution [22] popularly deployed in 4G and 5G networks [1], [4]. Let us commence by describing the conventional system without the learning network. In the CU of Fig. 1, the laser diodes (LDs) emit the unmodulated light into the mach-zenhder modulator (MZM) for the electro-to-optical conversion, where the RF signal conveying the quadrature amplitude modulation (QAM) symbols are carried by the light signal. Then, the outputs of the different MZMs in Fig. 1 are combined by the WDM multiplexer (WDM Mux), forming a WDM signal transmitted in the optical fiber from the CU to the RRH, as shown in Fig. 1. The WDM demultiplexer (WDM DeMux) separates the individual wavelengths, which are then converted back to the electronic signal after each photo-detector. Afterwards, the electronic bandpass filter (EBPF) and electronic amplifier (EA)

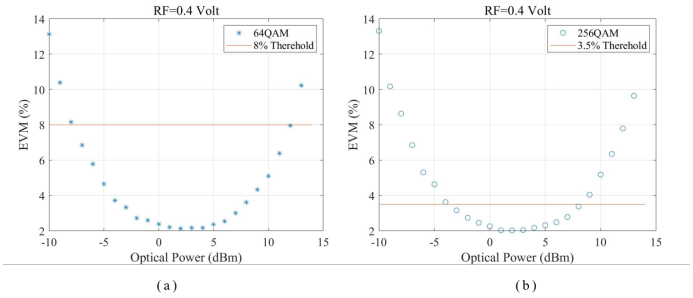

Fig. 2. EVM performance of the conventional system using 64-QAM and 256-QAM

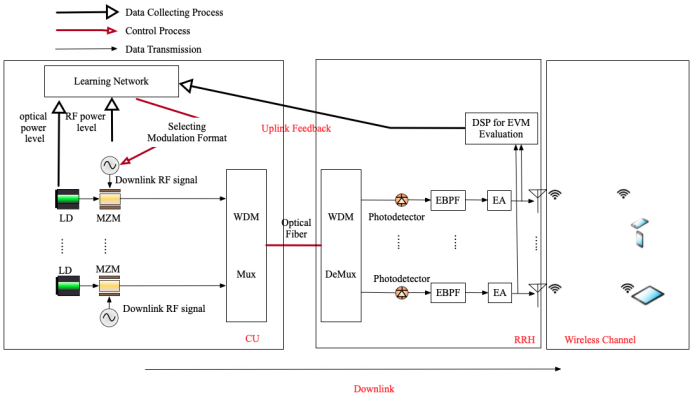

Fig. 3. A-RoF Fronthaul Downlink using logistic regression classification.

are responsible for filtering out the unwanted RF signal and the amplification before the RF signals are transmitted via the air by the antennas.

Then, in order to further elaborate the challenges of reduced-robustness and poor system performance, in the following, we present the system performance and the robustness of 64-QAM and 256-QAM using the parameters shown in Table I, which are in accordance with IEEE's 802.11ac standards [1]. Explicitly, the 64-QAM or 256-QAM signal is carried by the RF signal of 3GHz which is then optical-to-electronic (O/E) converted to the optical signal of 1550 nm wavelength over the standard single-mode fiber. Our simulation implements the carrier aggregation and OFDM with a sub-carrier spacing of 312.5 KHz, transmitting a wide-band signal of 800 MHz per wavelength.

Fig. 2 shows the dynamic range achieved when using 64-QAM and 256-QAM, which represents the laser power range meeting the 3GPP's EVM requirements, such as 8% for 64-QAM and 3.5% for 256-QAM. In both graphs of Fig. 2 and as shown in Table I, we transmit the signal at the same symbol rate (i.e. occupying the same bandwidth), where in Fig. 2a, the 64-QAM is implemented, while in Fig. 2b, the 256-QAM is explored. Thus, we have the bit rate ratio of $3:4$ between 64-QAM and 256-QAM due to their identical symbol rate, where the corresponding dynamic range ratio is $20:12$ as seen in Fig. 2. As a result, the 256-QAM has a higher data-rate but poorer dynamic range. Therefore, a good trade-off balancing the bit rate and the dynamic range should be investigated for boosting the system performance and robustness, which would benefit from the selection of the modulation formats.

The selection of 64 QAM versus 256 QAM is a classification problem, which has been investigated in the literauture [15], [16], while suffering from a high complexity. Hence, in the following, the machine learning algorithm using logistic regression model is invoked for reduced complexity [15], [23], in order to flexibly customize the A-RoF system in varied environments, thus improving the robustness, and adaptively selecting the modulation format.

III. Logistic Regression Model for A-RoF System Downlink

In this section, we present the learning aided A-RoF downlink system using logistic regression model, which is an example implementation of a learning algorithm. Logistic regression is a supervised learning method for predicting the optimal classes by training the historic data with a decision boundary. As mentioned in Section I, the major issues of A-RoF aided RAN system are its lack of robustness and poor system performance, which requires repeated system configuration customization for modulation format, optical power, RF signal etc. in order to have improved user experience. As a mature technique, logistic regression method features its flexibility, and simplicity in order to stabilize the system performance and improving the robustness, with the aid of adaptively selecting the modulation format according to the available power level.

Specifically, there are four phases for the logistic regression implementation in the A-RoF system, namely data collection, training phase, data validation and real-time data transmission. As shown in Fig. 3, a learning network for data training and modulation format selection is adopted in the CU of Fig. 3, whilst the data collection is performed in the RRH of Fig. 3 using "DSP for EVM evaluation" block. Specifically, the data collected by the "DSP for EVM evaluation" block is retrieved by the learning network for data training. Let us elaborate a little further, as shown in Fig. 3, a learning network is imposed in the CU, where the data from the "DSP for EVM evaluation" will be fed back to the learning network for determining the optimal modulation format based on the optical power level and RF power level. Then, the learning network will analyse the data based on the supervised learning using the logistic regression classification, where quasi-newton method, regularization and the feature normalization are used for enhancing the prediction accuracy. Finally, the trained model is built for testing the real-time data, where we can select the optimal RF power and optical power for enhanced channel capacity and system flexibility. In the following, we will elaborate on each phase.

A. Data Collection

According to the specific optical power and RF voltage, we collect the optimal class meeting the 3GPP's requirements as shown in Algorithm 1, which performs the method for the data collection in the "DSP for the EVM Evaluation" of the

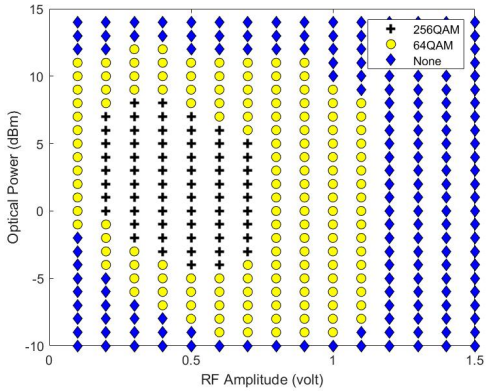

Fig. 4. Data Collection

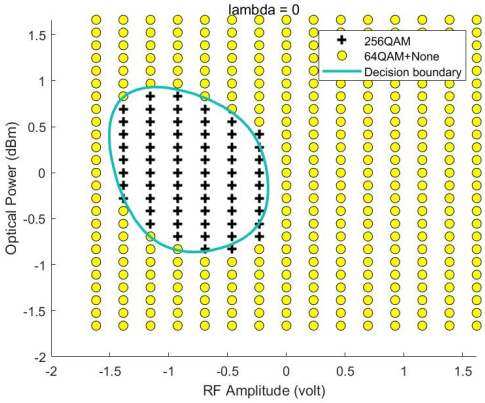

Fig. 5. Decision Boundary for 256QAM and 64QAM

RRH of Fig. 3, where the data collected will then be fed to the learning network of the CU.

Algorithm 1 Data Collection

1: Features including optical power and RF voltage
2: The CU collects the data point modulated by 64QAM or 256QAM
3: Data Collection
if EVM $\leq 3.5\%$ **then**
Select as 256QAM
elseif EVM $\leq 8\%$ **then**
Select as 64QAM
else
Select as Non
endif

The EVM threshold meeting the 3GPP's EVM requirement for 256 QAM and 64 QAM are 3.5% and 8% [1], [5], respectively. Therefore, by setting the range of the launched optical power and of the RF Voltage from -10 dBm to 15 dBm and from 0 V to 1.5 V, respectively, where the EVM is probably meeting the above threshold, and by invoking Algorithm 1, we collect the training data using MATLAB by simulating the A-RoF system described in Section II. Fig. 4 shows the training data, where the +, circle and diamond represent the 256 QAM, 64 QAM and other lower order modulation formats potentially providing better EVM performance, which we mark in the figure as "None".

These data will then be used as the training data for training the logistic regression algorithm in the CU of Fig. 3 for building the learning model based on logistic regression algorithm in the next section.

B. Training Phase

Fig. 4 shows the collected training data for the example simulation setup shown in Table I. In this section, by modelling a decision boundary using the logistic regression method, we can predict the optimal modulation formats to use at a specific power level. Assuming that $\mathbf{X} = \mathbf{\{x,y\}}$ is a matrix, where $\mathbf{x}$, $\mathbf{y}$ are the column vectors representing the optical power and RF Amplitude, respectively. To derive the decision boundary, the data is firstly normalized and then the Quasi-Newton algorithm is used for determining the optimal hypothesis function, where the logistic cost function are optimized [24].

The feature normalization aims to speed up the gradient descent algorithm by scaling the $\mathbf{X}$ with the standard deviation σ, resulting a new $\mathbf{X_{scaled}} = \{\mathbf{x_{scaled}}, \mathbf{y_{scaled}}\} = \{\frac{\mathbf{x-mean(x)}}{\sigma(\mathbf{x})}, \frac{\mathbf{y-mean(y)}}{\sigma(\mathbf{y})}\}$.

In this system, we attempt to find the optimal parameter θ of the hypothesis function $h_\theta(\mathbf{X_{scaled}})$ to minimize the objective function $J(\theta)$, where the problem is as [25], [26]

$$\min_{\theta} J(\theta), \tag{1}$$

where $J(\theta) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(\mathbf{X^{(i)}_{scaled}}) - \mathbf{z^{(i)}})^2 + \lambda\sum_{j=1}^{n}\theta_j^2)$, with m and n representing the number of training samples and the coefficient θ in the hypothesis function $h_\theta(\mathbf{X_{scaled}})$. The second term in the bracket is the regularizer to avoid overfitting [26]. The logistic regression cost function is as [26]

$$\begin{aligned}&\text{Cost}(h_\theta(\mathbf{X_{scaled}}) - \mathbf{z})\\ &= \mathbf{z}\log h_\theta(\mathbf{_{scaled}}) + (1-\mathbf{z})log(1-h_\theta(\mathbf{x_{scaled}})).\end{aligned} \tag{2}$$

Here, we use the binary classification to calculate the minimized $J(\theta)$, where $\mathbf{z}$ can be either 0 or 1. Note that the hypothesis function $h_\theta(\mathbf{X_{scaled}})$ is a sigmoid function as

$$h_\theta(\mathbf{X_{scaled}}) = \frac{1}{1+e^{\theta^{\mathbf{T}}\mathbf{X_{scaled}}}}. \tag{3}$$

Then, we use the MATLAB's built-in function **fminunc** to find the minimum of Problem (1) using the quasi-newton algorithm, where the best parameters of θ can be used for plotting the decision boundary as shown in Figs. 5 and 6 to predict the best modulation format. Specifically, we set the regularization parameters λ and j to be 0 and 28 for obtaining the optimal decision boundary. Then, we validate our model accuracy by feeding the original data as discussed in the following.

As analysed in Section III-B, the decision boundaries of Figs. 5 and 6 are used to predict the best modulation format

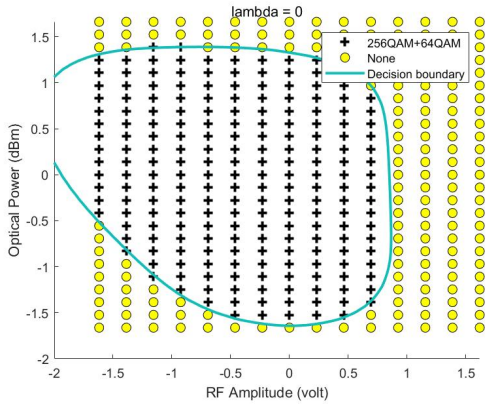

Fig. 6. Decision Boundary for QAM and None

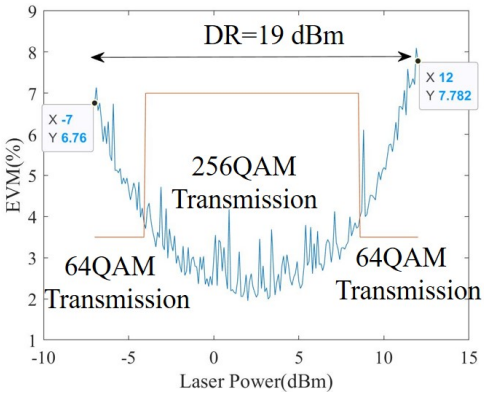

Fig. 7. The dynamic range of the proposed system

using the regularization parameter λ of 0 to avoid the underfitting, where Fig. 5 aims to classify differentiate 256 QAM and 64 QAM using $h_{\theta_1}(\mathbf{X_{scaled}})$, while Fig. 6 show how to classify using QAM or other schemes exploiting $h_{\theta_2}(\mathbf{X_{scaled}})$. Algorithm 2 based on $h_{\theta_1}(\mathbf{X_{scaled}})$ and $h_{\theta_2}(\mathbf{X_{scaled}})$ predicts the optimal modulation format.

Then, by feeding the original data as collected in Fig. 4, we compare our predicted modulation format with the data shown in Fig. 4, obtaining a training accuracy of 96.266667%. Next section presents the real-time demonstration by inputting random data into our proposed system.

Algorithm 2 Data Validation

if $h_{\theta_1}(\mathbf{X_{scaled}}) \geq 0.5\%$ **And** $h_{\theta_2}(\mathbf{X_{scaled}}) \leq 0.5\%$ **then**
  Select 256QAM
elseif $h_{\theta_1}(\mathbf{X_{scaled}}) \leq 0.5\%$ **And** $h_{\theta_2}(\mathbf{X_{scaled}}) \geq 0.5\%$ **then**
  Select 64QAM
else
  Select Others (i.e. those below 64-QAM)
endif

IV. Simulation Results and Analysis

In this section, we input the random data while setting the optical power ranging from -10 dBm to 15 dBm[1]. Specifically, the input optical power level and the input RF power level are evaluated by the learning network of the CU of Fig. 3. The logistic regression model performed in the learning network of the CU of Fig. 3 will select the optimal modulation format[2] used for the downlink RF signal based on Algorithm 2. Note that Logistic regression has a low computational complexity, which makes it practical for low latency applications [18], [19], [26], while the conventional closed-form solution using normal equation [25] will be very slow when data is very large [18], [19].

Then, the modulated downlink RF signal is E/O converted by the MZM of Fig. 3, and then the RF signal is transmitted by the A-RoF system to the RRH for a more-robust and enhanced-data-rate wireless transmission. As shown in Fig. 7, a real-time transmission selecting 64 QAM and 256 QAM is demonstrated. The red-line of Fig. 7 illustrates the power range for selecting 256-QAM or 64-QAM, with the upper line representing the 256-QAM transmission and the bottom line the 64-QAM transmission. Note that the results in this section are generated using the parameters in Table I.

As analysed in Fig. 7, 64 QAM is selected when the optical power ranges from -7 dBm to -4 dBm and from 9 dBm to 13 dBm, while the system selects 256 QAM when the power spans from -4 dBm to 9 dBm. Furthermore, we can conclude from Fig. 7, the dynamic range has been improved from the 12 dBm of only using 256 QAM for transmission as shown in the right graph of Fig. 2 to 19 dBm as shown in Fig. 7. Thus, the proposed system avoids the re-customization of the system parameters, such as the RF power, optical power, waveform, etc., since we can adaptively select the best modulation format based on the on-line training data, which reflects the context, to maintain the system reliability.

Meanwhile, the spectral efficiency is enhanced by a factor of $\frac{4}{3}$ when the power ranges from -4 dBm and 9 dBm, thanks to the transmission of 256 QAM. Note that when the optical power is lower than -4 dBm and higher than 13 dBm, neither 64 QAM nor 256 QAM is capable of meeting the 3GPP's requirement, so other modulation formats below 64-QAM should be considered.

Fig. 8 shows the probability density function (PDF) of the modulation formats along side the optical power range, where we can predict the best modulation format based on the optical power given. Hence, by on-line updating the training data, we can obtain updated optimal logistic regression model, adaptively predicting the optimal modulation formats, improving the system robustness and channel capacity, and

[1]For the power range lower than -10 dBm and those larger than 15 dBm, the EVM perfomance exceeds the 3GPP's EVM requirement due to the poor SNR and due to the fiber-nonliearity, respectively [1], which would be automatically regarded as "None"

[2]In this paper, we evaluate the optimal optical power whilst fixing the RF power as an example to verify that the learning aided logistic regression method can be used in our design. The RF power can also be optimised using the same process.

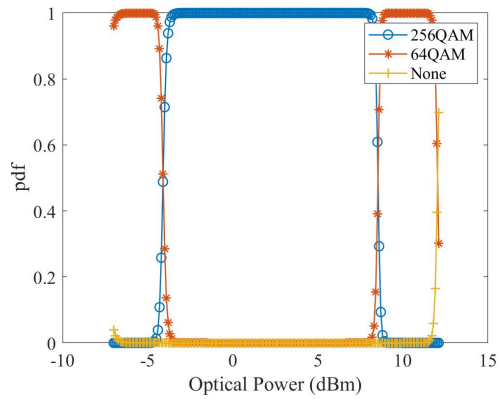

Fig. 8. Probability density function

increasing the dynamic range.

V. Conclusions

In this paper, inspired by the low-cost and bandwidth-saving of the A-RoF technique in the C-RAN, we proposed a learning aided A-RoF system using supervised learning. The less-robustness and poor-performance were addressed using the learning network implemented in the CU, where we invoked the logistic regression classification and improved the data rate by a factor of $\frac{4}{3}$ and dynamic range of $\frac{19}{13}$. Furthermore, we can adaptively select the modulation format relying on the training model, effectively removing the need for re-customization, which is seen as the major obstacle for the A-RoF's commercialization. Furthermore, the nonlinearity-induced poor dynamic range can be extended from 12 to 19 dBm thanks to the fast learning aided logistic regression method.

References

[1] Y. Li, F. Wang, M. El-Hajjar, and L. Hanzo, "Analog radio-over-fiber-aided optical-domain MIMO signal processing for high-performance low-cost radio access networks," *IEEE Communications Magazine*, vol. 59, no. 1, pp. 126–132, 2021.

[2] Y. Li, M. El-Hajjar, and L. Hanzo, "Joint space-time block-coding and beamforming for the multiuser radio over plastic fiber downlink," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 3, pp. 2781–2786, 2018.

[3] Y. Li, S. Ghafoor, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Analogue wireless beamforming exploiting the fiber-nonlinearity of radio over fiber-based C-RANs," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2802–2813, 2019.

[4] J. S. Wey, Y. Luo, and T. Pfeiffer, "5G wireless transport in a PON context: An overview," *IEEE Communications Standards Magazine*, vol. 4, no. 1, pp. 50–56, 2020.

[5] C. Consortium *et al.*, "eCPRI Specification V2. 0," May 2019.

[6] M. Tornatore, G.-K. Chang, and G. Ellinas, *Fiber-Wireless Convergence in Next-Generation Communication Networks*. Springer, 2017.

[7] D. Novak, R. B. Waterhouse, A. Nirmalathas, C. Lim, P. A. Gamage, T. R. Clark, M. L. Dennis, and J. A. Nanzer, "Radio-over-fiber technologies for emerging wireless systems," *IEEE Journal of Quantum Electronics*, vol. 52, no. 1, pp. 1–11, 2016.

[8] E. I. Ackerman, G. E. Betts, W. K. Burns, J. C. Campbell, C. H. Cox, N. Duan, J. L. Prince, M. D. Regan, and H. V. Roussell, "Signal-to-noise performance of two analog photonic links using different noise reduction techniques," in *2007 IEEE/MTT-S International Microwave Symposium*, pp. 51–54, 2007.

[9] F. S. Yang, M. E. Marhic, and L. G. Kazovsky, "Nonlinear crosstalk and two countermeasures in SCM-WDM optical communication systems," *Journal of Lightwave Technology*, vol. 18, no. 4, pp. 512–520, 2000.

[10] S. K. Korotky and R. M. de Ridder, "Dual parallel modulation schemes for low-distortion analog optical transmission," *IEEE Journal on Selected Areas in Communications*, vol. 8, no. 7, pp. 1377–1381, 1990.

[11] J. Marti, "Experimental reduction of chromatic dispersion effects in lightwave microwave/millimetre-wave transmissions using tapered linearly chirped fibre gratings," *Electronics Letters*, vol. 33, pp. 1170–1171(1), June 1997.

[12] X. N. Fernando and A. B. Sesay, "Adaptive asymmetric linearization of radio over fiber links for wireless access," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 6, pp. 1576–1586, 2002.

[13] H. Sotobayashi and K. Kitayama, "Cancellation of the signal fading for 60 GHz subcarrier multiplexed optical DSB signal transmission in nondispersion shifted fiber using midway optical phase conjugation," *Journal of Lightwave Technology*, vol. 17, no. 12, pp. 2488–2497, 1999.

[14] V. A. Thomas, M. El-Hajjar, and L. Hanzo, "Performance improvement and cost reduction techniques for radio over fiber communications," *IEEE Communications Surveys Tutorials*, vol. 17, pp. 627–670, Secondquarter 2015.

[15] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, 2017.

[16] F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An overview on application of machine learning techniques in optical networks," *IEEE Communications Surveys Tutorials*, vol. 21, no. 2, pp. 1383–1408, 2019.

[17] J. He, J. Lee, S. Kandeepan, and K. Wang, "Machine learning techniques in radio-over-fiber systems and networks," *Photonics*, vol. 7, no. 4, 2020.

[18] S. Ramkishore, P. Madhumitha, and P. Palanichamy, "Comparison of logistic regression and support vector machine for the classification of microstructure and interfacial defects in zircaloy-2," in *2014 International Conference on Soft Computing and Machine Intelligence*, pp. 130–134, 2014.

[19] O. Ergul and L. Gurel, "Iterative solution of the normal-equations form of the electric-field integral equation," in *2007 IEEE Antennas and Propagation Society International Symposium*, pp. 1857–1860, IEEE, 2007.

[20] J. Kim, M. Sung, S. Cho, Y. Won, B. Lim, S. Pyun, J. K. Lee, and J. H. Lee, "Mimo-supporting radio-over-fiber system and its application in mmwave-based indoor 5g mobile network," *Journal of Lightwave Technology*, pp. 1–1, 2019.

[21] M. El-Hajjar and L. Hanzo, "Multifunctional mimo systems: A combined diversity and multiplexing design perspective," *IEEE Wireless Communications*, vol. 17, no. 2, pp. 73–79, 2010.

[22] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud ran for mobile networks—a technology overview," *IEEE Communications Surveys Tutorials*, vol. 17, pp. 405–426, Firstquarter 2015.

[23] M. Kulin, T. Kazaz, E. De Poorter, and I. Moerman, "A survey on machine learning-based performance improvement of wireless networks: Phy, mac and network layer," *Electronics*, vol. 10, no. 3, p. 318, 2021.

[24] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, "Machine learning and the physical sciences," *Reviews of Modern Physics*, vol. 91, no. 4, p. 045002, 2019.

[25] A. Ng, "Advice for applying machine learning," in *Machine learning*, 2011.

[26] S. Dreiseitl and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: a methodology review," *Journal of biomedical informatics*, vol. 35, no. 5-6, pp. 352–359, 2002.