# A new LightGBM-based Equalizer enabled high-capacity PAM-4 and NRZ transmission in the 10-G class system

**Chenyu Wang (1), Xin Miao (2), Yu Xi (1), Meihua Bi (1,2), Weisheng Hu (2)**

(1) *School of Communication Engineering, Hangzhou Dianzi University, Xiasha Gaojiangyuan 2rd Street, Hangzhou, Zhejiang province, 310038, China. E-mail: bmhua@hdu.edu.cn*

(2) *State Key Laboratory of Advanced Optical Communication System and Networks, Shanghai Jiao Tong University, 800Dongchuan Rd, Shanghai, 200240, China. E-mail: wshu@sjtu.edu.cn*

***Abstract***—A powerful LightGBM-based equalizer is presented to improve the performance of bandwidth-limited IMDD system. Compared with the common 37-taps DFE, it can achieve the sensitivity by 2 dB at the HD-FEC limit.

***Keywords****—short reach LightGBM, DSP*

## I. Introduction

In recent years, with the rise of cloud computing, self-driving, virtual reality, and other future applications, both industry and business have higher requirements for optical access network and short distance communication [1]. Meanwhile, due to their huge number of end-users, these systems are extremely cost-sensitive [2]. Therefore, keeping both high capacity and low cost is the focus of these system designs, especially in the field of data center optical interconnection [3]. To achieve this target, a simple but efficient way is the high-rate data transmission with low-bandwidth optical devices. Besides, the high-speed optical transmission system with the low-bandwidth intensity modulation and direct detection (IM/DD) is also regarded as one of promising technology in short-distance optical transmission systems due to their low complexity and low cost [4]. However, for the high-speed systems with limited-band transceivers, the inter-code interference (ISI) induced limited bandwidth of devices, and cumulative dispersion of optical fibers result in a serious deterioration in system performance [5]. To solve this problem, many pre/post signal equalization methods have been proposed [6]. Among them, decision feedback equalizer (DFE) and feed-forward equalizer (FFE) are the mainstream digital signal processing technologies used to compensate for distortion and restore the signal. However, their large computational overhead and limited performance improvement result in low system transmission efficiency. Meanwhile, to further improve the equalization performance, more and more machine learning and neural network algorithms were also demonstrated [7]. However, traditional machine learning solutions often fail to achieve the desired results. Neural networks use complex structures to improve performance but at the same time increase computational overhead, making it difficult to implement in real-time systems within the low-cost short reach optical transmission systems.

To achieve the target of high capacity and low cost in the bandwidth-limited IM/DD system, we design a LightGBM-based equalization scheme. It is then applied to a system based on 25 Gb/s NRZ and 50 Gb/s PAM-4 in 10 GHz optical devices to manifest its feasibility. Moreover, we measure the performance based on FFE and DFE for comparative analysis. Results indicate that, compared to the 37-taps DFE, our method obtains 2.5 dB performance improvement in both 20km standard single-mode fiber (SSMF) and KP4-FEC threshold (BER=2.2E-4) back-to-back (B2B) transmission in the case of 25 Gb/s NRZ. For the 50 Gb/s PAM-4 proposed scheme can improve the sensitivity by 2 dB under the HD-FEC limit (BER=3.8E-3) compared to the common DFE with 37 taps.

## II. Principle of LightGBM scheme

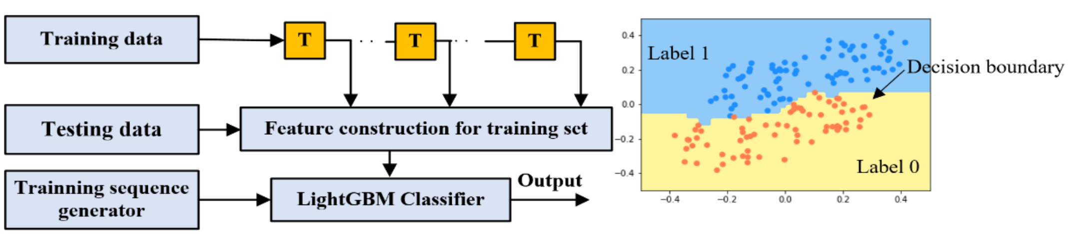


Fig. 1. Scheme diagram of our proposed LightGBM equalizer.

Figure 1 gives the main structure diagram of the LightGBM classification equalizer that we proposed. First, the training sequence is composed of partial signals passing through the receiving end of the system. The feature vector of the training sequence is constructed by the tap delayer, which can be described as,

$$\mathbf{X}_i = \left[x_{i-n}(d), x_{i-n+1}(d), \ldots x_i(d) \ldots, x_{i+n-1}(d), x_{i+n}(d)\right] \tag{1}$$

where $X_i$ represents the ***i***-th sampled data, and *d* represents the ordinal number of the feature column selected by down-sampling. n represents the number of previous samples and subsequent samples used for feature construction and also refers to the order of the feature. After the feature construction, the feature vector and label of the training data are used as the input of the LightGBM trainer.

Here, the LightGBM is used as an equalizer for obtaining the corresponding tap coefficient. As a framework, the LightGBM can implements the GBDT algorithm efficiently for classification [8]. Based on this idea, a simple classifier (decision trees) would be used to iteratively train to get the optimal equalization. Throughout these processes, the trained leaner can be finally achieved, which can be given as,

$$f(x) = f_0(x) + \sum_{t=1}^{T} \sum_{j=1}^{J} c_{tj} I\left(x \in R_{tj}\right) \tag{2}$$

where $f_0(x)$ is the initialized base learner, which can be understood as a constant. $T$ is the number of week learners, and $J$ is the index of the leaf nodes in each weak learner. $c_{tj}$ represents the value of leaf node $j$ in the t-th tree, $I$ is the index function, and the expression $I\ \{x \in R_{tj}\}$ indicates that if sample $x$ is in the leaf node $R_{tj}$, then return 1 otherwise return 0. In the equalization stage, valid data is input to the trained tree model to obtain the corresponding predicted value, and then this value is used as the input of the sigmoid function, which can be described as,

$$\sigma(x) = \frac{1}{1+\mathrm{e}^{-x}} \tag{3}$$

This function maps the calculated value to the interval of (0,1) to get the probability that the sample label is 1. Then the output is used to judge with the threshold of 0.5, and the judgment result is the final equalization result. At the same time, Fig. 1 also shows the visualization of the equalization effect of the LightGBM equalizer. The decision boundary obtained from the training sample classifies the sample points very well, which also confirms the equalization effect of the scheme.

## III. Experimental setup

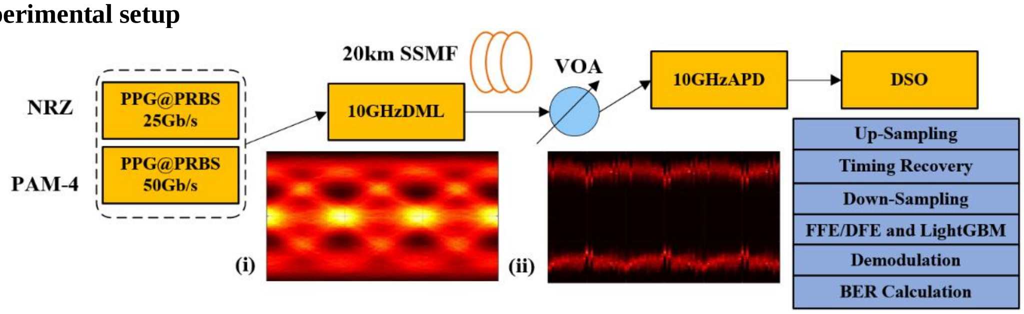


Fig. 2 Experimental setup of our scheme. Inserts: Eye diagram of (i) NRZ signals and (ii) after LightGBM equalization

Figure 2 gives the corresponding experiment configuration of this paper, which is utilized to manifest our scheme. The 10-Gb/s transceivers are employed in this system. At the transmitting end, the NRZ signals with 25Gb/s and PAM-4 signals with 50Gb/s as transmitted data are generated offline, whose length is $2^{15}$-1 pseudorandom binary sequence (PRBS-15) columns. These signals with 250,000 symbols length and the specifically-designed training sequence are uploaded into a pulse pattern generator (PPG) to generate the transmitted signals and then are utilized to drive a commercially direct modulation laser (DML) to achieve optical modulated signals. This DML has ~8 GHz modulated bandwidth and 1311-nm work wavelength and 10-dBm output optical power. Throughout 20 km SSMF transmission, the optical signals are transmitted into a variable optical attenuator (VOA) for further BER measure. At the receiving end, a ~ 7-GHz bandwidth (3-dB point) avalanche diode (APD) is utilized to detect the optical signals. The detected signals are then captured by an 80-GSa/s sampling rate oscilloscope (DSO, LeCroy SDA845Zi-A). At last, the sampled signals are injected into the offline DSP module, which is deployed by MATLAB includes up-sampling, timing recovery, down-sampling, equalization (including the common FFE/DFE and LightGBM) operations and, BER calculation.

## IV. Experimental results and discussion

To verify this scheme, the eye diagram of NRZ before and after our equalization is presented in Fig. 2(i) and (ii), respectively. Obviously, for our scheme, the eye becomes large which can achieve a good BER performance. Further, we also give a comparison analysis of DFE, FFE, and LightGBM with different taps at -24dBm received power, as depicted in Fig. 3(a). Within the converges range, our method is better than others even with the same number of features. And when the taps are set to 37(18 orders), LightGBM can improve the system performance by an order of magnitude.

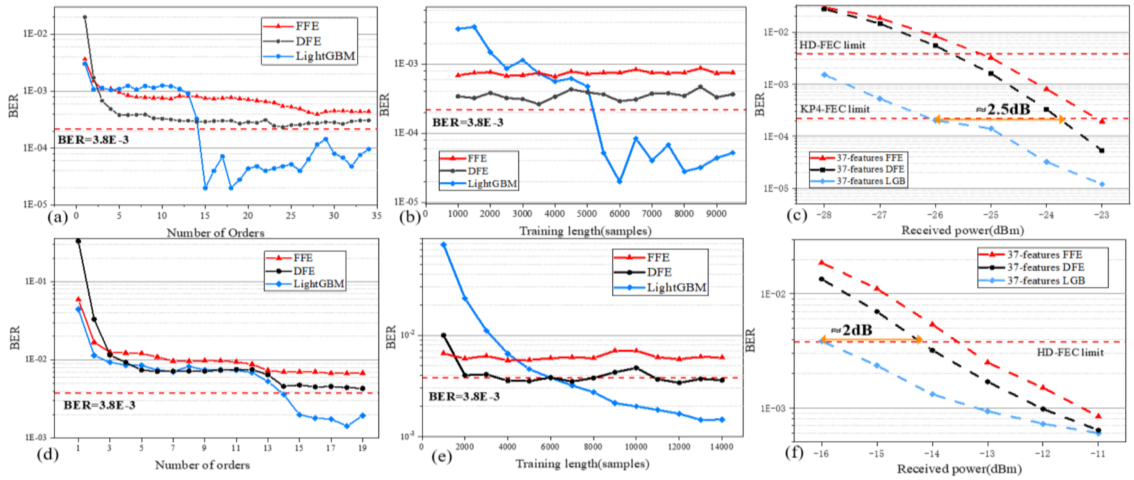


Fig.3 BER curves of NRZ with (a) different order numbers at -24dBm power. (b) different training length at -24dBm power. (c) different received power. BER curves of PAM-4 with (d) different order numbers at -14dBm power. (e) different training length at -14dBm power. (f)different received power.

Then, the convergence speed and the training sample length of the DFE, FFE, and LightGBM at -24dBm received power are given for further manifesting our scheme. As shown in Fig. 3(b). Here, the three schemes select the number of taps in their respective convergence states. It is easily got that, our solution requires more training samples to achieve the best results, by using the training length that greater than 5500 better performance than DFE and FFE. Besides, unlike DFE and FFE, our method also uses an iterative algorithm to train the model, but it can use an early-stop optimization method to reduce overfitting and reduce unnecessary computational overhead. Similar results can be achieved as for the PAM-4 signals which indicate in Fig. 3(d) and (e). And, for simplification, the corresponding description is neglected in this letter. Finally, we evaluate the BER performance for the B2B and 20km transmission. Here, for the sake of fairness, the parameter of the presented each equalizer have been configurated the optimal states of itself (DFE / FFE is the optimal value within 37 taps and its optimal training lengths, LightGBM is 37 taps and 6000/10000 training lengths respectively for NRZ and PAM-4). Also, LightGBM can get a considerable performance improvement. And comparing to the 37-taps DFE and FFE, LightGBM obtains 2.5dB sensitivity improvement @ BER = 2.2E-4 (KP4 FEC limit) for 25Gb/s NRZ and 2dB @ BER = 3.8E-3 (HD-FEC limit) for 50Gb/s PAM-4.

## V. Conclusion

We propose a new equalization method based on LightGBM for the IMDD-based bandwidth-limited optical system in this letter. This is verified to perform better in both NRZ and PAM-4 signals than the common DFE with the appropriate training length and taps. In future work, we will strive to reduce the training length used by the LightGBM scheme.

## Acknowledgment

This work was jointly supported by the Natural Science Foundation of Zhejiang Province [No. LY20F050004] and the Scientific research project of Zhejiang Provincial Department of Education [No. Y201942093, No. Y201942104].

## References

[1] P. Miguelez. What applications are driving higher capacity in access? Proc. Conf. Lasers Electro-Opt. Conf. (CLEO), pp. 1–3, Mar. 2018.

[2] J. Zhang et al. SOA Pre-Amplified 100 Gb/s/λ PAM-4 TDM-PON Downstream Transmission Using 10 Gbps O-Band Transmitters. Journal of Lightwave Technology, vol. 38, no. 2, pp. 185-193, 15 Jan.15, 2020.

[3] J. Chen, A. Tan, Z. Li, T. Xu, L. Liang, W. Chen, Y. Li, and Y. Song. 50-km C-Band Transmission of 50-Gb/s PAM4 Using 10-G EML and Complexity-Reduced Adaptive Equalization. IEEE Photonics J. 11(1), 1–10, 2019.

[4] S. Barthomeuf, F. Saliou, L. Anet Neto, G. Simon, F. Bourgart, P. Chanclou, and D. Erasme. Equalization and Interoperability Challenges in Next Generation Passive Optical Networks. Optical Fiber Communications Conference and Exhibition (OFC), (Optical Society of America), 2019.

[5] L. Xue, L. Yi, W. Hu, R. Lin and J. Chen. Optics-Simplified DSP for 50 Gb/s PON Downstream Transmission using 10 Gb/s Optical Devices. Journal of Lightwave Technology, vol. 38, no. 3, pp. 583-589, 1 Feb.1, 2020.

[6] Li D, Song H, Cheng W, et al. Low-complexity equalization scheme for suppressing FFE-enhanced in-band noise and ISI in 100 Gbps PAM4 optical IMDD system. Opt Lett. 45(9):2555-2558, 2020.

[7] S. Peng, H. Jiang, H. Wang, H. Alwageed. and Y. Zhou. Modulation Classification Based on Signal Constellation Diagrams and Deep Learning.IEEE T. Neur. Net. Lear. 30(3), 718–727, 2019.

[8] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q & Liu, T. Y. Lightgbm: A highly efficient gradient boosting decision tree. In Advances in neural information processing systems (pp. 3146-3154), 2017.