

Combining Efficient Probabilistic Shaping and Deep Neural Network to Mitigate Capacity Crunch in 5G Fronthaul

Qi Zhou^{1*}, Rui Zhang¹, You-Wei Chen¹, Shuyi Shen¹, Shang-Jen Su¹, Jeffrey Finkelstein²,
and Gee-Kung Chang¹

¹*School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30308, USA*

²*Cox Communications, Atlanta, GA 30328*

*qi.zhou@gatech.edu

Abstract: We experimentally demonstrate a capacity-approaching transmission in 5G fronthaul utilizing PS-PAM8 and DNN. An 80-Gb/s over 20-km SSMF transmission performance is realized with a beyond 7.3-dB gross gain over uniform PAM modulations with linear post-equalization.

OCIS codes: (060.2330) Fiber optics communications; (060.4080) Modulation.

1. Introduction

Stimulated by the exponential growth on the demand of data-thirsting services such as immersive virtual/augmented reality and super-resolution 4K/8K video, the bandwidth requirements in 5G radio access network have increased tremendously. 5G promises to provide capability to support such bandwidth-hungry applications by introducing new mm-wave frequencies. However, due to the short transmission range of mm-waves, new network paradigm has been recently adopted, wherein next generation mobile network (NGMN) is split into three major units: central unit (CU), distributed unit (DU) and remote radio unit (RRU), as illustrated in Fig. 1(a). The connection between CU and DU is typically referred to as the midhaul while the link between the DU and RU is known as the fronthaul [1-3]. Besides the architecture-wise innovation like flexible function-split based on different service types [4], the mobile fronthaul (MFH) also requires the deployment of high-capacity and low-cost fiber based physical infrastructures. Among those systems, intensity modulation/direct detection (IMDD) is a preferred scheme due to its low power consumption, small footprint and low cost. On the other hand, those features of the IMDD scheme bring in some drawbacks as well, including limited transmitting power and small channel bandwidth. PAM modulations like OOK and PAM4 are widely used in IMDD since they are simple and robust. However, the PAM modulation lacks sufficient flexibility in adjusting spectral efficiency, which makes it incapable to fully utilize the channel resources (i.e., bandwidth and SNR). Originated from coherent optical communications, probabilistic shaping (PS) has become a popular approach to reduce the gap in SNR between the capacity of the optical communication system and the Shannon limit [5]. Through shaping the occurrence probability of the constellation points in a modulation scheme, or in other words, increasing the possibility to transmit lower amplitude symbols, we can increase the minimum Euclidean distance among all the constellation points at a given average signal power as illustrated in Fig.2(b). Besides, the spectral efficiency of PS modulation can be continuously adjusted by varying the constellation probabilistic distribution. Nevertheless, the PS shaped signal is associated with a higher PAPR, which can easily exceed the small dynamic range of the low-cost IMDD based MFH link, resulting in severe nonlinear impairments. Inspired by the booming progress in artificial intelligence, neural network based nonlinear equalizer has found its applications in optical communication recently [6]. Moreover, neural networks with multiple hidden layers, i.e., deep neural network (DNN), demonstrate even better performance. In our previous work, we demonstrated a DNN decoder for enhanced multi-level signal recovery in MFH which outperformed the conventional Volterra nonlinear equalizer (VNLE) [7]. Although there are some concerns about the DNN complexity, the training time of the DNN to adapt dynamic channel environment can be significantly reduced through transfer learning once the initial training is complete. Moreover, the online querying of the trained DNN, i.e., nonlinear equalization and signal decoding, just involves multiple matrix multiplications with a low time complexity [8].

In this paper, for the first time, we efficiently take advantage of both PS to utilize the channel resources and DNN on nonlinear signal equalization/decoding, and experimentally demonstrate a capacity-approaching transmission in the 5G MFH. A proof-of-concept experiment is conducted to verify the efficacy of the proposed scheme with an 80-Gbps PS-PAM8 signal over 20-km standard single mode fiber (SSMF) transmission based on SD-FEC.

2. Operation principles and experimental setup

Fig.1(c) shows the experimental setup of the proposed scheme for NGMN. Firstly, a uniformly distributed binary bit sequence is randomly generated and input to a distribution matcher (DM) for PS. Here, the implemented DM is based on the constant composition distribution matcher (CCDM) [9]. The PS symbols output by the CCDM follow the Maxwell-Boltzmann distribution with flexible entropy through adjusting the constellation point occurrence probability

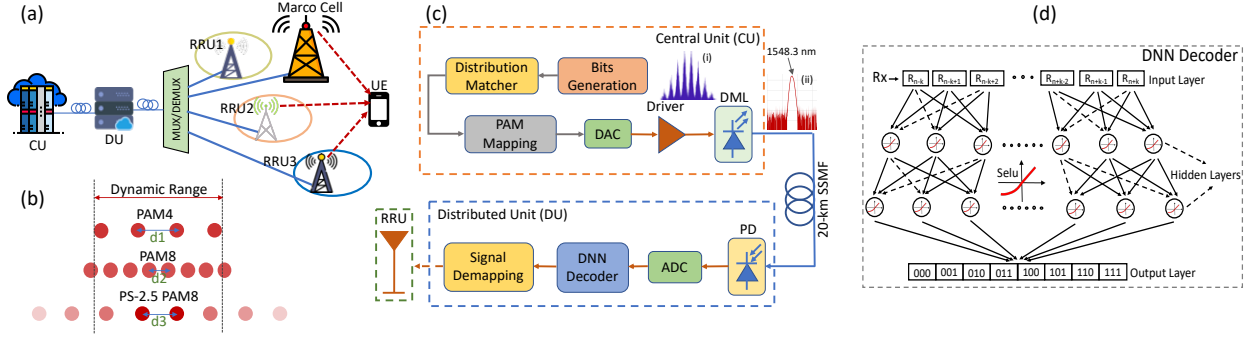


Fig. 1 (a) Illustration of NGMN architecture. (b) Illustration of PS-2.5 PAM8 with high PAPR; d_1 , d_2 , d_3 are minimum Euclidean distance for PAM4, PAM8, and PS-2.5 PAM8, respectively. (c) Experimental setup of the proposed transmission scheme, CU-DU-RRU; including inset (i) Histogram of the generated PS-2.5 PAM8 signal and (ii) Optical spectrum of the DML output. (d) Structure of the DNN decoder.

as given by: $P(a_i) = \exp(-\lambda a_i^2) / \sum_i (-\lambda a_i^2)$, where a_i is the amplitude of a constellation point, and λ is a fitting parameter. The shaped symbols are then PAM mapped and converted to an analog signal as shown by inset (i) through a 64-GSa/s arbitrary waveform generator (AWG, M8195A). The modulator driver is used to provide sufficient signal voltage swing to drive the DML (SCMT-100M11G), while the DML is used as the transmitter in the CU to convert the driving signal into the optical domain with a center wavelength at 1548.3 nm as illustrated in inset (ii). The optical signal propagates through a 20-km SSMF and detected by a PIN receiver (SCMR-100M11G) in the DU. The detected signal is digitized using an oscilloscope with 80-GSa/s sampling rate (DSOZ254A) for the following digital signal processing. As shown in Fig.1(b), though the PS-PAM8 signal with an information rate of 2.5 bits/symbol has a larger minimum Euclidean distance than the uniform PAM8 modulation at a certain average signal power, it also introduces a higher PAPR. The peak of the signal can easily exceed the small dynamic range of the low-cost electrical and optical components in the MFH and make it vulnerable to nonlinear impairments. In this case, commonly used linear equalization schemes (i.e., least mean square algorithm (LMS)) will not suffice. Nevertheless, DNN is proven to be able to eliminate the nonlinear impairments efficiently because of its superior modeling capability owing to the multi-layer architecture and nonlinear activation function. Moreover, implementing the DNN into the MFH is more practical now because of the explosive improvements in efficient nonlinear multi-variable optimization algorithms and the low-cost parallel computing hardware (i.e., graphic processing unit (GPU) and tensor processing unit (TPU)). Fig.1(d) shows the structure of the implemented DNN decoder. We take the input symbol with its 80 previous and 80 subsequent symbols as the input layer. The DNN consists of 2 hidden layers with 1024 neurons at each layer. There are 8 neurons at the output layer, each one corresponds to a constellation point of the PAM8 modulation. We use the Selu function as the activation function at the hidden layers which can effectively eliminate the gradient vanishing and gradient exploding problems that commonly result in DNN's incapability of convergence. The activation function at the output layer is Softmax given by: $\sigma(z)_j = e^{z_j} / \sum_{k=1}^K e^{z_k}$. The Softmax function is employed to calculate the probability of each DNN output values, while categorical cross-entropy defined as: $J(w) = -1/N \sum_{n=1}^N [y_n \log \hat{y}_n + (1 - y_n) \log (1 - \hat{y}_n)]$, is implemented for loss calculation. In the combination of the Softmax function and the categorical cross-entropy, the calculation of the partial derivative for back-propagation is more concise. Besides, we utilize the Adamax algorithm to train the DNN, which can efficiently avoid the local optima. Another severe problem associated with DNN is overfitting. We alleviate it through integrating dropout layers in the DNN as shown by the dashed blue arrows in Fig.1(d), which deactivates 20% of neurons during each training epoch. We also divide the received symbols into 3 sets, namely, training set, validation set, and testing set with a ratio of 0.6, 0.1, and 0.3, respectively. The validation set is used to measure the validation accuracy over the whole training process. When the validation set accuracy is not improving over 100 epochs, we will stop the training to avoid severe overfitting.

3. Experimental results

To verify the feasibility of the proposed scheme, we conduct a proof-of-concept experiment. The frequency response of the end-to-end system configured as Fig.1(c) is measured by a 20-GHz vector network analyzer as shown by the red curve of Fig.2(a). The measured S21 response is in between the input to the modulator driver and the PIN receiver output. As observed, the S21 response is relatively flat below 16 GHz, while it degrades dramatically beyond that frequency. For comparison, we also plot the spectra of PS-PAM8, PAM8 and PAM4 with a bit rate of 80 Gb/s in Fig.2(a) as indicated by blue, purple, and green curve, respectively. Due to the low spectral efficiency of PAM4 modulation, the spectrum of PAM4 signal is far beyond the bandwidth of the system. However, since we can flexibly set the entropy of PS-PAM8 signal, the PS-PAM spectrum closely matches the system frequency response, which maximizes the bandwidth utilization. As for the PAM8 signal, though it's within the bandwidth limit of the system,

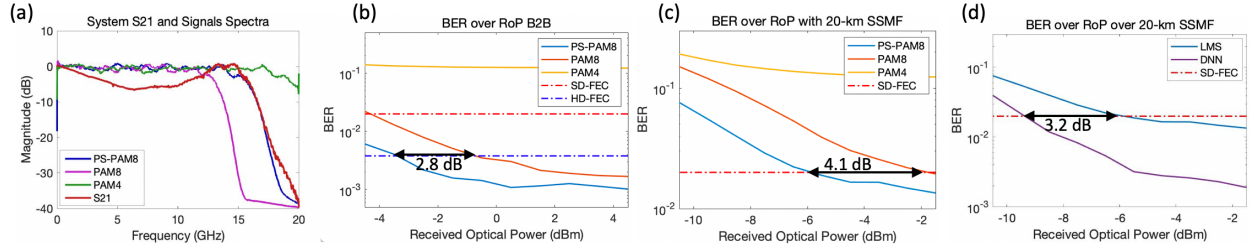


Fig. 2 (a) Measured system S21 response and signals spectra. BER over received optical power (RoP), (b) at B2B case, (c) over 20-km SSMF, (d) with and without DNN decoder.

the Euclidean distance between adjacent constellations is too small to reach acceptable BER performance because of the limited SNR in the low-cost IMDD system. To set up a baseline in reference to the discussed modulation schemes, we firstly measure the optical back-to-back (BtB) transmission performance of 80-Gb/s signals based on IMDD with an optimized LMS post-equalization. During the measurement, the received optical power (RoP) is scanned from -4.5 dBm to 4.5 dBm with 1-dB increment. The measured BER curves for PS-PAM8, PAM8 and PAM4 are shown in Fig.2(b) as blue, red, and yellow solid curves, respectively. The SD-FEC and HD-FEC thresholds are also plotted as reference lines. For the PAM4 modulation, the BER cannot go below either of the thresholds for all the measured RoP values due to the insufficient system bandwidth. In the PAM8 modulation case, the BER becomes lower as the RoP increases and passes the HD-FEC threshold at -0.6-dBm RoP. However, the PS-PAM8 achieves a -3.4-dBm receiver sensitivity and demonstrates a prominent 2.8-dB gain over PAM8 modulation at the HD-FEC threshold. In addition, the BER curves after 20-km SSMF transmission are also measured as shown in Fig.2(c). The PS-PAM8 demonstrates a remarkable 4.1-dB gain compared with PAM8 at SD-FEC threshold. As discussed in the former section, the higher power efficiency of the PS signal is at the cost of increasing the PAPR. In the low-cost MFH system, the high PAPR can easily exceed the limited dynamic range of the transmission link and resulting strong nonlinearity. The DNN decoder is an efficient method for mitigating nonlinear impairments. To verify the performance, we randomly generate a 100k symbol length PS-PAM8 signal in the CU. At the DU, after resampling and resynchronization, we take the first 60% of the received symbols as training set, the following 10% as validation set, and the remaining 30% is test set for the DNN. We monitor the validation-set accuracy while training the DNN, once validation-set accuracy stops improving over 100 epochs, the algorithm will stop the training automatically and save the most current DNN model parameters. The trained DNN decoder will then decode the test-set data. Fig.2(d) demonstrates an experimental verification on the DNN's performance. The blue curve is the baseline BER performance over RoP with the LMS post-equalization, while the red curve indicates the performance of the DNN decoder. At the SD-FEC threshold, the DNN decoder demonstrates a noteworthy extra 3.2-dB gain over the baseline, which proves the DNN's efficacy on eliminating the probabilistic shaping's drawback of nonlinear impairment, and on significantly improving the overall transmission performance from the CU to DU.

4. Conclusions

We propose and experimentally demonstrate a capacity-approaching transmission in 5G MFH based on PS-PAM8 modulation and DNN decoder. The PS-PAM8 signal is power efficient and is flexible in information rate to adapt the channel conditions. However, probabilistic shaping introduces a higher PAPR and make the PS signal vulnerable to the limited dynamic range of the hardware components in 5G MFH resulting in severe nonlinear impairments. We subsequently implement a DNN decoder for nonlinear compensation to significantly improve the overall transmission performance from the CU to DU with a 3.2-dB extra gain at SD-FEC threshold. Adding up the 4.1-dB gain from PS-PAM8, a 7.3-dB gross gain is realized comparing with conventional uniform PAM modulations with linear post-equalization. Our scheme offers a promising solution to mitigate capacity crunch in 5G MFH.

5. References

- [1] T. Pfeiffer, J. Opt. Commun. Netw.7, B38 (2015).
- [2] G.-K. Chang, et al. "Key Fiber Wireless Integrated Radio Access Technologies for 5G and Beyond," in 2019 24th OptoElectronics and Communications Conference (OECC), (IEEE, 2019), pp. 1-3.
- [3] P. Chancelou, et al., J. Opt. Commun. Netw.10, A1 (2018).
- [4] Y. M. Alfadhli, et al. "Real-time Demonstration of Adaptive Functional Split in 5G Flexible Mobile Fronthaul Networks," in Optical Fiber Communication Conference, (Optical Society of America, 2018), pp. Th2A-48.
- [5] J. Cho, et al., J. Light. Technol. 37, 1590 (2019).
- [6] F. Musumeci, et al., IEEE Commun. Surv. & Tutorials 21, 1383 (2018).
- [7] Q. Zhou, et al., IEEE Photonics Technol. Lett. 30, 1511 (2018).
- [8] J. Zhang, et al., Opt. letters 44, 4243 (2019).
- [9] P. Schulte, et al., IEEE Transactions on Inf. Theory 62, 430 (2015).