

# ANN-based Optimization of Probabilistic and Geometric Shaping for Flexible Rate 50G and beyond PON

Shuang Yao<sup>1,3</sup>, Amitkumar Mahadevan<sup>1</sup>, Yannick Lefevre<sup>2</sup>, Noriaki Kaneda<sup>1</sup>, Vincent Houtsma<sup>1</sup>,  
and Doutje van Veen<sup>1</sup>

1. Nokia Bell Labs, 600 Mountain Ave, Murray Hill, NJ, USA

2. Nokia Bell Labs, Copernicuslaan 50, Antwerp, Belgium

3. Georgia Institute of Technology, School of Electrical and Computer Engineering, Atlanta, GA, USA  
syao65@gatech.edu, amitkumar.mahadevan@nokia-bell-labs.com

**Abstract:** Joint probabilistic and geometric shaping is considered for flexible PON. Optimal modulation is found through ANN, which generalizes to various ROPs and fiber lengths while taking dispersion, limited bandwidth, and receiver-side DSP into account. © 2022 The Author(s)

## 1. Introduction

Passive optical networks (PON) employ point-to-multipoint (PTMP) architecture and serve as a cost-effective solution for services such as fiber-to-the-home (FTTH) and mobile x-haul. In downstream direction, an optical line terminal (OLT) broadcasts to optical network units (ONUs) over a shared feeder fiber using a fixed modulation format and line rate that are standardized (e.g. 50G NRZ in ITU-T G.hsp [1]) to ensure that an ONU with the worst case channel impairments (loss and/or dispersion) can still meet performance requirements. This results in inefficient resource utilization since a larger percentage of ONUs experience less severe impairments in practice. Therefore, flexible rate PON where the information rate to each ONU is optimized, is an attractive solution to improve overall system throughput and reduce latency [2].

Flexibility can be achieved in several ways. For example, the modulation format can be switched from codeword to codeword, as in time domain hybrid modulation [3,4]. It can also be implemented by means of constellation shaping, which can provide flexibility as well as shaping gain. In our previous works [2,5], we investigated the feasibility of a flexible rate PON by adopting joint probabilistic shaping (PS) and geometric shaping (GS) through both numerical simulations and experiments. The information rate was adjusted by applying PS+GS modulation that maximizes the generalized mutual information (GMI). The optimal distributions were found at each received optical power (ROP), without establishing a relationship between the optimal PS+GS modulation and the various channel conditions typically observed in a PTMP system. This work extends the flexible PON simulation testbed developed in [5] by taking fiber dispersion and receiver side equalization into account, and also deploys an artificial neural network (ANN) which allows for the learning of optimized distributions under various channel conditions.

An ANN is a promising candidate to solve a wide range of problems in optical communications [6] due to its versatility. In principle, an ANN with one hidden layer and adequate number of hidden neurons can approximate any continuous function [7]. Thus, it is expected that it can also approximate the relationship between the various channel conditions and the optimized PS+GS modulation parameters. Compared with previous works on ANN-based distribution optimization [8,9], we adopt a channel model that characterizes higher rate PONs (50G and beyond), wherein peak and average power constraints are applied, and distortions including avalanche photodiode (APD) noise and inter-symbol-interference (ISI) arising from the use of bandwidth limited components at the receiver and chromatic dispersion from up to 20 km of fiber transmission are considered. The receiver employs a 25G class APD and a 15-tap feedforward equalizer (FFE) that is comparable to the transmitter and dispersion eye closure (TDEC) reference receiver adopted in the G.hsp standard [1]. By utilizing the developed channel model, the ANN is capable of generalizing to both ROP and fiber distance. According to simulation results, continuous tuning of the line rate can be achieved at all the test cases where ROP and fiber length vary. Further, a GMI improvement of above 0.1 bits/symbol compared with uniform PAM signals is attained.

## 2. Working principles

Fig. 1(a) illustrates the deployed ANN and the system model. The ANN as well as the entire simulation testbed described below is implemented in Tensorflow. In this paper, symmetric PS+GS PAM4 is adopted to realize flexible rate PON, which offers a smooth transition to lower modulation orders. As shown in Fig. 1(b), the symmetric constraint on the probability distribution allows for the insertion of the uniformly distributed forward error correction (FEC) parity bits at the most significant bit (MSB). The input features are ROP and fiber length of the ONU. Considering that the optimal modulation is a piecewise-defined function of ROP, a one-hot encoding vector of ROP is also included at the input layer. Specifically, the ROP is divided into three regions: 1) low ROP: when ROP is  $\leq -28.5$  dBm; 2) medium ROP: when ROP is within the range  $(-28.5, -22.5)$  dBm; 3) high ROP: when ROP is  $\geq -22.5$

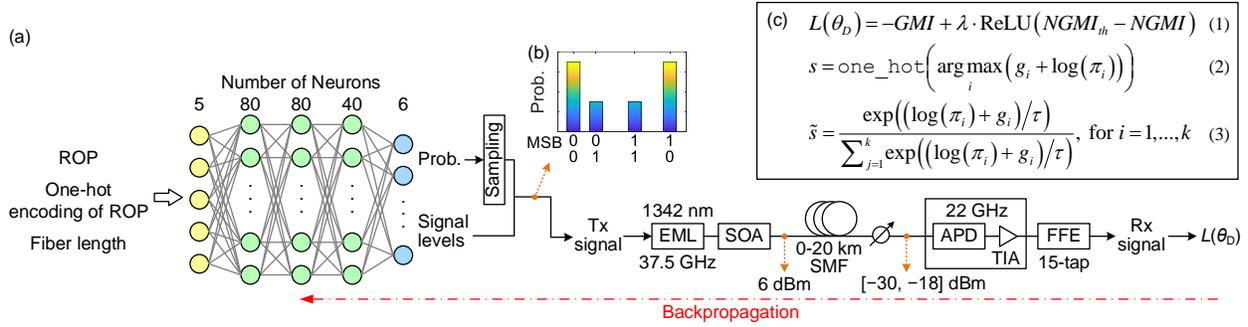


Fig. 1. (a) The deployed ANN and system model; (b) symmetric PS+GS PAM4 modulation; (c) Eq. (1)-(3).

dBm. The regions are encoded as  $[1, 0, 0]$ ,  $[0, 1, 0]$ , and  $[0, 0, 1]$ , respectively. There are three hidden layers with 80, 80, and 40 neurons, respectively. Rectified linear unit (ReLU) activation functions are applied at the hidden layers. The output layer has 6 neurons, two of them corresponding to the probabilities of the two symbols with lowest power and the other four corresponding to the signal levels.

To account for chromatic dispersion, receiver-side bandwidth limitation and equalization, Monte Carlo simulations are conducted during training to calculate GMI. For each forward propagation, a sampling process generates PS samples according to the probabilities defined by the ANN output. A PS+GS PAM signal is then generated by applying GS to these PS samples. It is delivered through the channel and processed by receiver-side DSP. An externally modulated laser (EML) with bandwidth of 37.5 GHz, chirp of 0.5 and wavelength of 1342 nm is utilized as proposed for 50 GPON in ITU-T G.hsp. The peak power constraint is applied at the EML, where the outermost signal levels of PAM4 (i.e.,  $-3$  and  $3$ ) are mapped to 1 mW and 4 mW, respectively. The average power constraint is applied at the SOA, which compensates for power loss due to shaping and ensures an average launch power of 6 dBm into the fiber. The single mode fiber (SMF) has a dispersion coefficient of 3.686 ps/(nm·km), which is the worst case for 1342 nm [10]. A 22-GHz bandwidth receiver consists of an APD and a transimpedance amplifier (TIA). Signal-amplified spontaneous emission (ASE) beat noise, thermal noise and shot noise are added onto the signal. The power of the shot noise is assumed to be proportional to the average ROP of the signal. A 15-tap FFE at the receiver compensates for the bandwidth limitations and incurred fiber dispersion. The loss function is calculated using the signal after the DSP and the transmitted signal. It is a function of the ANN weights and biases  $\theta_D$  and is defined in Eq. (1) in Fig. 1(c). The first term is GMI, which is the achievable information rate assuming ideal FEC encoding and decoding [11]. The GMI is calculated based on a Gaussian approximation of the received signal after DSP, with the Gaussian statistics estimated at the receiver for each transmitted signal level. To ensure parity from a systematic FEC code can be transmitted, normalized GMI (NGMI) threshold is considered, and a second term that captures the penalty related with violation of NGMI requirement is added to the loss function.  $\lambda$  is the penalty coefficient, which is chosen as 40 when fiber length is  $\leq 5$  km and as 80 otherwise.  $NGMI_{th}$  is the NGMI threshold and is set as 0.5. When NGMI is higher than  $NGMI_{th}$ , the second term becomes 0 and thus the loss function only contains negative GMI. When the NGMI requirement is not satisfied, the loss function is dominated by the second term, prohibiting solutions having a NGMI higher than the threshold.

After calculating the loss function,  $\theta_D$  are updated with the Adam algorithm, which performs stochastic gradient descent (SGD) using gradients computed with backpropagation. However, the sampling process is not differentiable. To solve this issue, we adopt Straight-Through Gumbel Estimator [12]. Suppose PAM- $k$  symbols have probabilities  $\pi_1, \dots, \pi_k$ . In the forward propagation, the discrete PAM modulation is preserved by using Eq. (2) in Fig. 1(c), where  $s$  is a  $1 \times k$  vector that is the one-hot encoding of the PAM symbol, and  $g_i$  are independent and identically distributed (i.i.d.) random variables drawn from standard Gumbel distribution. The distribution of  $s$  is the same with the ANN output. In the backpropagation, softmax function is utilized to approximate the non-differentiable argmax operator using Eq. (3) in Fig. 1(c).  $\tau$  is a positive parameter that controls the approximation error of  $\tilde{s}$ . When  $\tau \rightarrow 0$ ,  $\tilde{s} = s$ . And when  $\tau \rightarrow +\infty$ ,  $\tilde{s} = [1/k, \dots, 1/k]$ .

### 3. ANN Evaluation and Results

During the training phase, ROP and fiber length are independently sampled from  $U(-30, -18)$  dBm and  $U(0, 20)$  km, respectively. The training is conducted with 2000 combinations of ROPs and fiber lengths. The test dataset comprises 525 datapoints constructed by sweeping ROPs from  $-30$  to  $-18$  dBm with a step size of 0.5 dBm and sweeping fiber lengths from 0 to 20 km with a step size of 1 km. The evaluation is also performed with Monte Carlo simulations, where at each test datapoint,  $2^{18}$  symbols are transmitted through the testbed to estimate the GMI. Fig. 2(a) shows GMI versus ROP in the case of 10-km SMF transmission. When ROP is lower than  $-28$  dBm, PS+GS PAM4 has

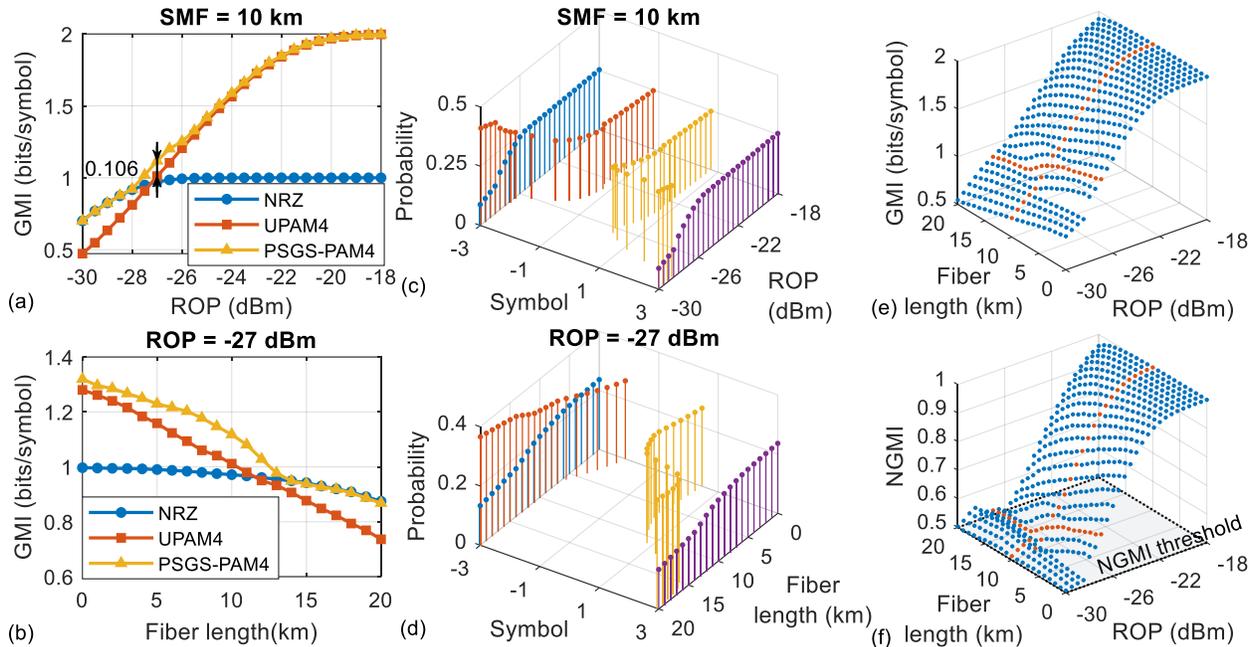


Fig. 2. (a) GMI versus ROP when fiber length is 10 km; (b) GMI versus fiber length when ROP is  $-27$  dBm; (c) PS+GS modulation versus ROP when fiber length is 10 km; (d) PS+GS modulation versus fiber length when ROP is  $-27$  dBm; (e)&(f) GMI and NGMI at all the test datapoints.

similar performance with uniform NRZ. When ROP is higher than  $-24$  dBm, there is negligible difference between GMI of PS+GS PAM4 and that of uniform PAM4. While for ROP larger than  $-28$  dBm and smaller than  $-24$  dBm, PS+GS PAM4 presents higher GMI and the improvement is 0.106 bits/symbol at the ROP of  $-27$  dBm. Fig. 2(b) plots the GMI versus fiber length with a fixed ROP of  $-27$  dBm. Again, PS+GS PAM4 outperforms both uniform NRZ and PAM4 over a wide range of fiber lengths. Fig. 2(c) and (d) plot the PS+GS modulation at the fixed fiber length of 10 km and at the fixed ROP of  $-27$  dBm, respectively. For a 10-km fiber, it is close to uniform NRZ at low ROPs and gradually converges to uniform PAM4 as ROP increases. Similarly, at a fixed ROP, as the dispersion increases, the distance between the inner two symbols increases and their probabilities become higher. At  $-27$ -dBm ROP with a 10-km fiber, the ANN outputs a constellation that balances the levels and probabilities providing the GMI improvement shown in Fig. 2(a).

Fig. 2(e) and (f) summarizes GMI and NGMI at all the test datapoints. The cases of 10-km SMF and  $-27$ -dBm ROP are highlighted with red. It can be clearly seen that continuous tuning of GMI can be achieved in all the cases, while meeting the NGMI threshold of 0.5 at the same time.

#### 4. Conclusions

In this paper, we consider symmetric PS+GS modulation to realize flexible rate PON and train an ANN to learn distributions that adjust the information rate to the channel conditions. By including ROP and fiber length as input features, the ANN can find optimized modulation settings in a PTMP system with various loss and chromatic dispersion values. With Monte Carlo simulations, the impact of bandwidth limitations, fiber dispersion, and equalization with an FFE are considered during optimization. Continuous tuning of GMI can be achieved while demonstrating GMI improvement of above 0.1 bits/symbol than uniform PAM signals.

#### 5. References

- [1] ITU-T G.9804.3, "50-Gigabit-capable passive optical networks (50G-PON): Physical media dependent (PMD) layer specification," 2021.
- [2] N. Kaneda *et al.*, "First Experimental Demonstration of Flexible Rate PON Beyond 100Gb/s with Probabilistic and Geometric Shaping," *OFC '21*, F2H.2.
- [3] N. Kaneda *et al.*, "DSP for 50G/100G hybrid modulated TDM-PON," *ECOC '20*, Tu2J-6.
- [4] R. Borkowski *et al.*, "Operator Trial of 100 Gbit/s FLCSS-PON Prototype with Probabilistic Shaping and Soft-Input FEC," *ECOC '21*, We4F.3.
- [5] R. Zhang *et al.*, "Probabilistic and geometric shaping for next-generation 100G flexible PON," *ECOC '20*, Tu1J-6.
- [6] V. Houtsuma *et al.*, "92 and 50 Gbps TDM-PON using Neural Network Enabled Receiver Equalization Specialized for PON," *OFC '19*, M2B.6.
- [7] K. Hornik *et al.*, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, pp. 359–366, 1989.
- [8] M. Stark *et al.*, "Joint Learning of Geometric and Probabilistic Constellation Shaping," *Globecom Workshops*, 2019.
- [9] F. A. Aoudia and J. Hoydis, "Joint Learning of Probabilistic and Geometric Shaping for Coded Modulation Systems," *GLOBECOM*, 2020.
- [10] ITU-T G.652, "Characteristics of a single-mode optical fibre and cable", 2016.
- [11] J. Cho and P. J. Winzer, "Probabilistic Constellation Shaping for Optical Fiber Communications," *J. Lightw. Technol.*, vol. 37, no. 6, 2019.
- [12] E. Jang *et al.*, "Categorical reparameterization with gumbel-softmax," *Int. Conf. Learn. Representations*, 2017.

This work was partially supported by the Flemish Government funding agency VLAIO through the SPIC project (HBC.2020.2197).