An Optimization Method for Probabilistic Constellation Shaping in Peak-Power Constraint Systems in the Presence of Peak Enhancement Effects

Basak Ozaydin, Xi Chen, and Di Che*
Nokia Bell Labs, Murray Hill, NJ 07974, United States
**di.che@nokia-bell-labs.com*

Abstract: We propose a generic method to optimize the probabilistic distributions for a peak-power constraint system with arbitrary peak enhancement effects. The technique is useful for developing flexible-rate optical transceivers in links without optical amplifiers. © 2024 The Author(s)

1. Introduction

Probabilistic constellation shaping (PCS) has been widely deployed in commercial coherent optical transceivers as a capacity-approaching technique with the capability of fine rate adaptation to accommodate a variety of fiber channel conditions. In amplified coherent transmissions, the fiber launch power is limited by nonlinear Shannon limit, leading to an average-power-constrained (APC) channel. For an APC channel, it is known that the optimal PCS distribution is the Maxwell-Boltzmann (MB) distribution [1]. In contrast, the optimum PCS distribution in optical-amplifierless systems is not as straightforward. Without optical amplifiers, an optical transmitter is subjected to a peak power constraint (PPC) imposed by the electro-to-optical (E/O) modulator [2]. Some efforts have been made to find suitable PCS distributions under the PPC [3-6]. However, previous works usually ignored the fact that the peak of an analog signal that drives the E/O modulator is a function of both the static RF channel as well as the original constellation design, referred to as the "peak-to-average-power ratio enhancement" (PAPR-E) phenomenon as detailed in [2]. Such enhancement can happen anywhere from the generation of digital baseband symbols till the analog waveform reaches the E/O modulator. The mechanisms causing such enhancement include digital signal processing (*e.g.*, Nyquist pulse shaping, digital pre-emphasis, etc.), component bandwidth limits (*e.g.*, DAC, RF driver, etc.), and so on. The ignorance of PAPR-E can result in a potentially invalid design of the constellation. The complication of the impact of PAPR-E on PCS design is two-fold. First, the amount of PAPR-E changes from one transmitter to another. This means two transmitters with different components or DSP cannot share the same optimum PCS design. Second, even with the same transmitter, the amount of PAPR-E is variable and depends on the chosen baseband constellation. For instance, the peak enhancement is usually stronger for a uniform PAM signal than for a PCS PAM signal. Because it is not trivial to analytically characterize the relation between the baseband signal peak and the peak of the analog signal that drives the E/O modulator [2], it is difficult to optimize the PCS distribution for a practical PPC system. Several attempts were made to find a PCS distribution. One way is brute-force searching shown in [3] for a 4-PAM system. However, the brute-force method tends to have unacceptable complexity when the PAM order goes higher. The second method is searching for distributions among a set of distribution families [7] (like MB [2,7], reverse MB [2,3,7], and exponential [7] distributions); however, such a limited searching space is not guaranteed to contain the optimum distribution. In this paper, we propose a new low-complexity optimization scheme to find the capacity-achieving PCS distribution for PPC channels in the presence of *arbitrary* peak enhancement. The framework is applicable to any PPC system, like an intensity-modulation direct-detection (IM-DD) or a coherent one without optical amplifiers. Without loss of generality, we verify the scheme in a 200-Gb/s class IM-DD PAM system with both simulation and experiment.

2. Optimization model and procedures

The signals in an M-PAM constellation come from a symmetric and bipolar alphabet of cardinality M (the baseband constellation is treated as bipolar in an optical-ampliferless system, see [2]), and are distributed according to the distribution $p_X(x_i) = p_i$ that satisfies $p_X(x_i) = p_X(x_{M+1-i})$. We assume probabilistic shaping for which the levels in the PAM alphabet $\mathcal{X}$ are equally spaced, *i.e.*, $\mathcal{X} = \{-\Delta(M-1), -\Delta(M-3), \ldots, \Delta(M-1)\} = \Delta\{-(M-1), \ldots, M-1\} = \Delta\mathcal{X}_p$, where Δ is the scaling factor that determines the difference between adjacent levels of $\mathcal{X}$. We use a random variable X to represent the baseband PAM signal and a random variable Y as the baseband received signal. To generate the drive signal for the E/O modulator, X is up-sampled, filtered, and converted into an analog signal through a DAC, and then amplified by an RF amplifier. As mentioned, the peak power of the E/O drive signal depends on i) the chosen M-PAM constellation and distribution, ii) the DSP applied at the transmitter side (*e.g.*, pulse shaping, pre-equalization, etc.), and iii) the response of transmitter components (*e.g.*, DAC, RF amplifier). We model the peak power of the E/O drive signal with a function $f(p_X, \Delta)$ that maps a given distribution over the M-PAM constellation, p_X, and a scaling factor, Δ, to the peak power in $\mathbb{R}^+$. The optimization process aims to find a M-PAM signal that maximizes the mutual information (MI) between the baseband modulated signal X, and the received signal Y under the PPC. Thus, the optimization problem to determine the optimal constellation can be formulated as

$$C = max_{p\_X,\Delta}\, I_{p_X,\Delta}(X;Y)\,, where\ f(p_X,\Delta) = A \qquad (1)$$

where A is the PPC imposed by the E/O mapper. While (1) may appear to be a straightforward optimization problem, it should be noted that function $f(p_X,\Delta)$ is not easy to be analytically characterized. This is because even if the transfer functions of all the components on the transmitter side are known, the amount of PAPR-E depends on the digital baseband constellation and PCS distribution as mentioned. Prior works usually simplified the optimization process by ignoring PAPR-E or treating it as a constant for all modulation formats which is typically not true.

We propose a two-step approach to solve the optimization problem. The first step is to characterize $f(p_X,\Delta)$. We sample random distributions from $f(p_X,\Delta)$ and apply curve fitting on these samples to build a numerical model specifically for the transmitter of interest. Without loss of generality, we set $\Delta = 1$ in this step. For each randomly sampled distribution p, we create a symbol sequence (*i.e.*, L M-PAM symbols) and send it through the system under test. The 'system' can either be the real-world system, or a digital emulation of the system (*e.g.*, transfer function or frequency response of the system). The peak of the symbol sequence after passing through the system is measured and labeled as a training sample. The fitting method can be chosen according to the channel model and computation capabilities (*e.g.*, polynomials, or neural network (NN) fitting), and we choose NN fitting in this work. For instance, to characterize a 6-PAM ($\mathcal{X} = \pm1, \pm3, \pm5$) signal, two probabilities need to be specified, *i.e.*, $p_1 = P(X=1)$, $p_2 = P(X=3)$ (and $p_5 = 0.5 - p_1 - p_3$). In Fig. 1a, the NN predictions (blue) follows the actual observations very well. We then solve $f(p_X,\Delta)$ to express Δ as a function of p_X, i.e. $\hat{\Delta}(p_X)$. Consequently, the objective, $I_{p_X,\Delta}(X;Y)$, can be written as a function solely by the constellation distribution p_X. As a result, we obtain an unconstrained optimization problem over the set of probability distributions, *i.e.*, $C = \max_p I_p(X;Y)$, which can then be easily solved by common gradient ascend algorithms with multiple random initializations.

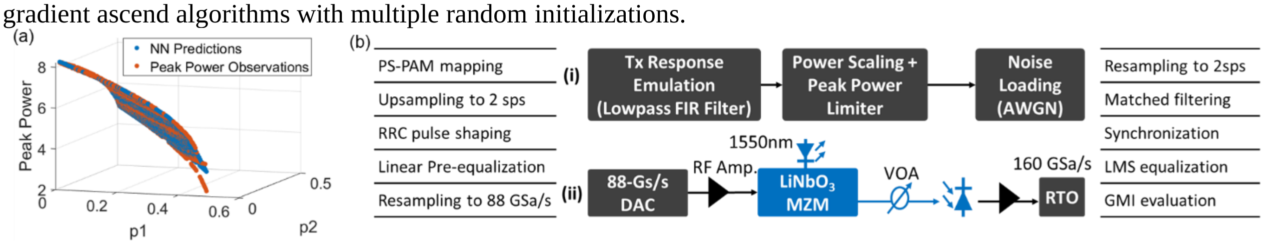

Fig. 1. (a) Curve fitting in comparison with the real peak observations (for 6-PAM, RRC 0.01); (b) simulation and experimental setup.

3. Results (from simulation and experiment) and discussion

We perform both simulations and experiment to verify the proposed framework using the IM-DD setup in Fig. 1b. Due to the complicated nature of PAPR-E, besides our approach, the only other known method so far that is bound to find the optimal constellation is the brute force approach. Therefore, we apply the brute force method as the ground truth for comparison. In our setup, the transmitter consists of an 88-GS/s D/A converter (DAC) with about 35-GHz analog bandwidth, a 55-GHz RF amplifier, and a $LiNbO_3$ Mach-Zehnder modulator (MZM). We use a variable optical attenuator (VOA) to adjust the received optical power, which correspondingly changes the channel peak signal-to-noise ratio (PSNR). As explained in [2], for PPC systems, PSNR (instead of SNR) should be the criterion for characterizing the *channel* quality. The signal is detected by a 70-GHz PIN photodiode (PD) whose output is amplified by a 60-GHz amplifier and sampled by a 62-GHz real-time oscilloscope (RTO) at 160 GSa/s. For the simulation, the transmitter is emulated by a lowpass finite impulse response (FIR) filter and a power scaling function to match the signal peak to a pre-defined PPC, and the PSNR is adjusted by adding various amounts of additive white Gaussian noise. As the peak is a statistical variable after PAPR-E, we define the peak as the waveform peak after 0.1% clipping.

First, we present simulations for the optimization of 8-PAM constellations. Fig. 2 displays the simulation results for root-raised cosine (RRC) filters with roll-off factors 0.01 and 0.2 to emulate two transmitters with strong and weak

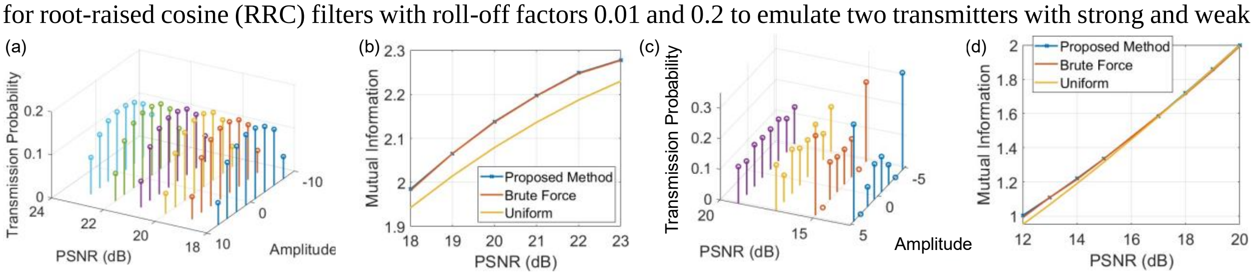

Fig. 2. 8-PAM simulation results for two transmitters with PAPR-E. (a-b) A strong PAPR-E case emulated by an RRC roll-off factor of 0.01 plus a linear pre-equalization, and (c-d) a weak PAPR-E case emulated by an RRC roll-off factor of 0.2.

PAPR-E respectively. We also add a linear pre-equalization M filter (max. 10 dB at the highest frequency) for the case of RRC 0.01 for an even higher PAPR-E. For RRC 0.01, we choose a PSNR range from 18 dB to 23 dB; while for RRC 0.2, we extend the PSNRs down to 12 dB, due to the weaker PAPR-E. We used a fully connected NN with a single hidden layer of 150 nodes to fit $f(p_X, \Delta)$. To train this neural network, we sampled only 200 points from $f(p_X, \Delta)$. Fig. 2(a,c) show the constellations obtained via our optimization. Fig. 2(b,d) show the MI for these optimized constellations in comparison with the brute force search. We also highlight the performance of a uniform 8-PAM signaling to show the PCS gain of our proposed method.

With the simulation results, we can make some interesting observations. We first focus on the strong PAPR-E case in Fig. 2 (a-b), where our method finds constellations with MIs (Fig.2b) that overlap with the ground truth (brute force method). The distribution of the constellations is close to MB as shown in Fig.2a. This confirms the predictions in [2]. In other words, MB distribution is a close-to-optimum choice given strong PAPR-E. This also coincides with experiments in some other literature where the symbol rate is being heavily pushed to the bandwidth limit of the transmitter therefore a strong PAPR-E presents [7,8]. In contrast, for the weak PAPR-E case in Fig. 2(c-d), the optimum distribution shows an interesting different trend. Theoretically, systems with weak PAPR-E would be close to an ideal PPC system, and the optimum distribution should be close to a uniform one at a high PSNR region according to [2]. This is verified in Fig. 2c high PSNRs (*e.g.*, 19 dB) At low PSNR region, the optimum distribution is similar to a reverse MB distribution. This is shown in demonstrations from [3,6]. The reverse MB achieves a gain in PPC systems mainly because the two edge PAM levels (with only one neighbor level) suffer less from wrong symbol decisions. Because of this, the gain is only obvious when PSNR is low, and quickly vanishes when PSNR is high [2]. What is even more interesting is that the distributions at low PSNRs in Fig. 2c do not fully follow the reverse MB distribution, but rather are a "double-cup" shape, serving as evidence that search of distributions within a certain group (*e.g.*, in reverse MB) may not be the best for all transmitters. In general, the observations in our simulations provide a comprehensive explanation of why previously published experiments yield different optimal distributions.

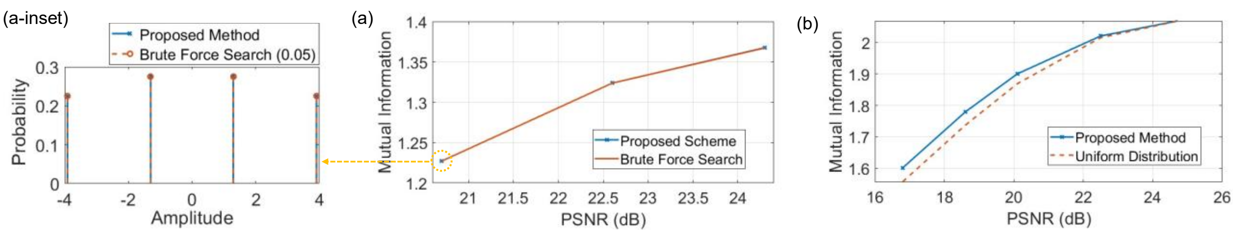

Fig. 3. 80-GBaud (RRC 0.01) PAM experimental results. MI as a function of PSNR for (a) 4-PAM (b) 8-PAM.

To verify the proposed scheme in the real world, we compare our proposed method with the brute force method in an 80-GBaud PAM-4 experiment (RRC 0.01), with the results shown in Fig. 3a. Both the MI curves in Fig. 3a and the example distributions (with a PSNR of 20.7 dB) in Fig. 3 (a-inset) show similar results for the two methods. This clearly indicates our proposed method indeed finds the global optimal distributions. Meanwhile, our gradient-ascend-based approach avoids the very high complexity of the brute-force method which increases exponentially with the PAM order. We also tried our scheme in an 80-GBaud PAM-8 system (RRC 0.01). Due to the prohibitively high complexity of the brute-force search for PAM-8, we only compare our method with uniform distributions. With a strong PAPR-E due to RRC filtering, PCS shows a shaping gain over the uniform signaling as expected, similar to what we found in Fig. 2b. Note that due to a weaker PAPR-E in the experiment, the PCS gain in Fig. 3b is less than that in Fig. 2b.

4. Conclusions

We have demonstrated a novel low-complexity constellation optimization method for PPC channels compatible to arbitrary PAPR-E effects. We showed that our constellation designs are as good as the distributions found with the brute force search method with much reduced computational complexity. The findings also reveal why previous literature demonstrated different optimal distributions in PPC experiments for the first time.

References

[1] F. Buchali et al., "Rate adaptation and reach increase by probabilistically shaped 64-QAM…," J. Lightw. Technol. 34, 1599 (2016).
[2] D. Che et al., "Does probabilistic constellation shaping benefit IM-DD systems without optical amplifier?" J. Lightw. Technol. 39, 4997 (2021).
[3] T. Wiegart et al., "Probabilistically shaped 4-PAM for short-reach IM/DD links with a peak power …," J. Lightw. Technol. 39, 400 (2021).
[4] S. Yao et al., "ANN-based optimization of probabilistic and geometric shaping for flexible rate 50G and beyond PON," OFC'2022, W3G.2.
[5] D. Kim et al., "Capacity-achieving symbol distributions for directly modulated laser and direct detection …," Opt. Express 31, 12609 (2023).
[6] B. M. Oliveira et al., "Revisiting probabilistic constellation shaping in unamplified coherent optical links," OFC'2023, Th3E.1.
[7] M. S.-B. Hossain et al., "Probabilistic shaping for high-speed unamplified IM/DD systems …," J. Lightw. Technol. 41, 5373 (2023).
[8] J. Zhang et al., "Demonstration of 260-Gb/s single-lane EML-based PS-PAM-8 IM/DD for datacenter interconnects," OFC'2019, W4I.4.