Flexible Coherent Communication System With Adaptable SNR and Laser Phase Noise Tolerance for Probabilistically Shaped QAM

Shuang Yao, You-Wei Chen, Rui Zhang, Xizi Tang, Qi Zhou, Shang-Jen Su, Shuyi Shen, Yahya Alfadhli, and Gee-Kung Chang, *Fellow, IEEE*

***Abstract*—Probabilistic-shaped quadrature amplitude modulation (PS-QAM) based on Maxwell-Boltzmann (MB) distribution has been extensively studied in recent years in the SNR limited coherent communication system to approach the Shannon capacity. However, MB distribution is not the optimal distribution when the transmission system suffers from serious laser phase noise. To improve the laser phase noise tolerance of MB distribution, a novel Maxwell-Boltzmann/Angular-distance-directed (MB/ADD) distribution is proposed in this article. The benefits of PS-QAM exist in the case of limited SNR, as well as large laser phase noise. The proposed distribution is verified through simulations and experiments. In our simulation analysis, the ADD part enables the reduction of SNR penalty when laser phase noise increases. The results are validated by experimental data, in which both digitally generated and distributed-feed-back (DFB) laser generated phase noise are employed to evaluate laser phase noise tolerance. In the scenario of large phase noise, MB/ADD shaping can achieve lower pre-FEC BER and higher GMI compared with MB shaping, without inserting pilot symbols. Moreover, the improvement of applying MB/ADD distribution can be obtained no matter whether BPS or V&V based algorithm is adopted for carrier phase estimation (CPE). Different shaping parameters are also tested, where the capability to agilely adapt to various channel conditions of the MB/ADD distribution is shown. And the proposed MB/ADD shaping can be applied with similar complexity with MB shaping, if the Kullback–Leibler (KL) divergence between the generated and the expected 2-D QAM distribution is required to be lower than 10^{-3}.**

***Index Terms*—Coherent communication, laser phase noise, probabilistic shaping.**

I. Introduction

COHERENT optical communication system has improved receiver sensitivity and higher spectrum efficiency (SE) as compared to direct detection system. The chromatic dispersion from fiber transmission can be compensated by digital signal processing (DSP), eliminating the complicated dispersion management. Coherent transceivers based on standard quadrature amplitude modulation (QAM) formats have played a central role in the long-haul [1] and high data rate optical networks [2].

However, there exists a gap to Shannon capacity for standard QAM formats [3]–[5], which prevents the full exploitation of the available channel resources. Constellation shaping, namely geometric shaping (GS) [6]–[8] and probabilistic shaping (PS) [3]–[5], have been proposed to solve this issue. In the GS, constellation points are no longer restricted to the rectilinear grid of standard QAM formats and the location of these points are rearranged to maximize achievable information rate (AIR). In spite of the shaping gain brought by GS, it often requires digital-to-analog converter (DAC) with higher effective number of bits (ENOB) [9] and more complicated DSP [5], as a result of the non-equidistant constellation points. PS, on the other hand, still preserves the square QAM formats and overcome the gap to Shannon capacity by shaping the occurrence of each point. When Maxwell-Boltzmann (MB) distribution is adopted, the entropy can be maximized under transmission energy constraint and up to 1.53 dB shaping gain can be achieved for the additive white Gaussian noise (AWGN) channel [10]. PS-QAM based on MB distribution has been extensively studied in recent years and record SEs have been achieved [11].

Despite the benefits brought by MB distribution, in coherent optical systems, there are various impairments apart from AWGN, such as laser phase noise, fiber nonlinearity, etc. And the shaping gain of MB distribution cannot be guaranteed under those circumstances. In [12], 256-QAM with uniform distribution presented higher optical signal-to-noise-ratio (OSNR) tolerance and AIR than PS-1024QAM with MB distribution when laser linewidth is 40 kHz and the pilot ratio is less than 5%. To fully exploit the advantages of MB distribution, lasers with narrower linewidth are mandatory, the high cost of which poses a great challenge on the massive deployment of flexible coherent transceivers.

To improve the adaptability and flexibility of the coherent optical systems in view of SNR and laser phase noise limitation, in this paper we propose a novel distribution by combining the MB distribution with an angular-distance-directed (ADD) distribution, namely an MB/ADD distribution. In the case of the system suffering from small laser phase noise and thus can

Manuscript received April 4, 2020; revised June 21, 2020; accepted July 16, 2020. Date of publication July 21, 2020; date of current version November 16, 2020. This work was supported in part by a grant from Industry/University Cooperative Research program of National Science Foundation for Center of Fiber Wireless Integration and Networking (FiWIN) for Heterogeneous Mobile Communications under Contract 1821819. *(Corresponding author: Shuang Yao.)*

The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30308 USA (e-mail: syao65@gatech.edu; yu-wei.chen@ece.gatech.edu; ruizhangece@gatech.edu; txizi3@gatech.edu; qi.zhou@gatech.edu; taiwanjen@gatech.edu; ssyzoe@gatech.edu; yalfadhli@gatech.edu; geekung.chang@ece.gatech.edu).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2020.3011076

be approximated as an AWGN channel, MB/ADD distribution will be similar to the conventional MB distribution. On the other hand, if the coherent transceiver employs low-cost large linewidth lasers, for instance when lasers used in the passive optical networks (PONs) are adopted for coherent communication [13], the ADD distribution will be assigned more weights in the MB/ADD distribution. Thus, the shaping gain from PS can be observed not only in the cases of limited SNR, but also in the case of large laser phase noise. Moreover, the trade-off between the SNR tolerance and the laser phase noise tolerance can be optimized by choosing proper shaping parameters. In this way, the flexibility of the conventional coherent communication system based on PS-QAM is assured by adding laser phase noise adaptability. The performance of the proposed MB/ADD distribution is verified through both simulations and experiments. In the simulation, the improvement is evaluated by the required SNR. In the experiments, the improvement is evaluated by pre-forward error-correction (pre-FEC) bit error rate (BER) and generalized mutual information (GMI). The received signals are tested under digitally generated phase noise, as well as distributed-feedback (DFB) laser generated phase noise. Results show that, in the region with large phase noise, the pre-FEC BER can be improved by one order of magnitude, compared with the traditional MB shaping. The flexibility is also demonstrated, where the performance of different combinations of shaping parameters is measured and discussed.

The rest of this paper is structed as follows. In Section II, MB/ADD distribution is introduced. In Section III, simulation results are presented. Experimental demonstration is described in Section IV, with results discussed in Section V. Finally, Section VI concludes the paper.

II. Principles of Operation

At a coherent receiver, assuming that perfect clock recovery, channel equalization and carrier frequency offset (CFO) can be achieved, the received signals can be written as:

$$y_k = x_k e^{j\varphi_k} + n_k \tag{1}$$

where x_k is the transmitted symbol, φ_k is the laser phase noise and n_k is the AWGN. If we only consider white phase noise, it can be modeled as a Wiener process:

$$\varphi_k = \sum_{i=-\infty}^{k} \varepsilon_i \tag{2}$$

where ε_i are independently identically distributed (i.i.d.) Gaussian random variables with zero mean and variance $\sigma_p^2 = 2\pi(\Delta f \cdot T_s)$. T_s is the symbol period. Δf reflects the combined effect of phase noise from transmitter laser and local oscillator (LO). If both lasers have Lorentz lineshape, Δf is the sum of their 3-dB linewidth [14]. Commonly used carrier phase estimation algorithms (CPE) include blind phase search (BPS) [15] and Viterbi and Viterbi (V&V) algorithm [16]. For 4-QAM (i.e. quadrature phase shift keying, QPSK) signals, CPE can be easily fulfilled by V&V algorithm and they also present high laser phase noise tolerance. However, the SE of 4-QAM signals is only 2 bits/Hz and in this paper we'll restrict our discussion to higher order QAM signals.

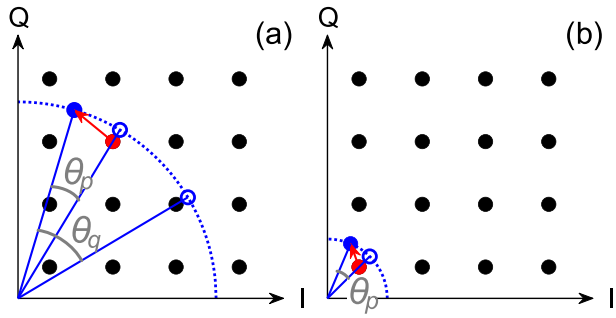

Fig. 1. Ambiguity issue of the BPS algorithm: (a) transmitted symbol is $3+5\boldsymbol{i}$; (b) transmitted symbol is $1+1\boldsymbol{i}$. (Red filled point: transmitted symbol; blue filled point: received symbol; blue circles: $\boldsymbol{y_k e^{-j\theta_p}}$ and $\boldsymbol{y_k e^{-j\theta_q}}$, where $\boldsymbol{\theta_p}$ and $\boldsymbol{\theta_q}$ minimizes Eq. (4)).

A. BPS

As described in [15], the BPS rotates the received signals by B test phases,

$$\theta_b = \frac{b}{B} \cdot \frac{\pi}{2}, b = 0, 1, \ldots, B-1 \tag{3}$$

The rotated symbols are sent into a decision module and the squared Euclidean distances between the rotated symbols and the decided symbols are calculated:

$$d_{k,b} = \left| y_k e^{-j\theta_b} - \left\lfloor y_k e^{-j\theta_b} \right\rceil_D \right|^2 \tag{4}$$

Test phases with smaller distance $d_{k,b}$ are regarded as better approximation of the carrier phase noise. Usually, $2N+1$ consecutive symbols are viewed as a block and their $d_{k,b}$ are summed up to get a distance metric $D_{k,b}$. θ_b that produces the minimum $D_{k,b}$ is assumed to be the phase noise, i.e.

$$\hat{\varphi}_k = \arg\min_{\theta_b} D_{k,b} = \arg\min_{\theta_b} \sum_{n=-N}^{N} d_{k-n,b} \tag{5}$$

The introduction of the block processing has a two-fold effect: i) it can mitigate the impact of AWGN on the phase noise estimation; ii) it can partly resolve the ambiguity issue in searching for the optimal test phase using Eq. (4). Assume that B is large enough so that we do not need to consider the resolution issue. Eq. (4) will be minimized when $y_k e^{-j\theta_b}$ has the same angle with $y_k e^{-j\theta_b}{}_D$. However, such θ_b is not unique as angular distances between points with the same magnitude can be less than $\pi/2$. As an example, Fig. 1(a) illustrates the case of 64-QAM. The red filled point represents the transmitted symbol and the blue filled point represents the received symbol, distorted by both AWGN and phase noise according to Eq. (1). In this example, there are two minimizers to the Eq. (4), denoted as θ_p and θ_q, with $y_k e^{-j\theta_p}$ and $y_k e^{-j\theta_q}$ illustrated by the blue circles. In the case of $N = 0$, such ambiguity will cause an error in the phase noise estimation.

Adopting a positive N can partly mitigate this issue, since it is highly unlikely that both θ_p and θ_q still minimize Eq. (5). However, this operation comes at the expense of reduced capability to track rapid phase changes, as the phase noise is assumed unchanged within the block.

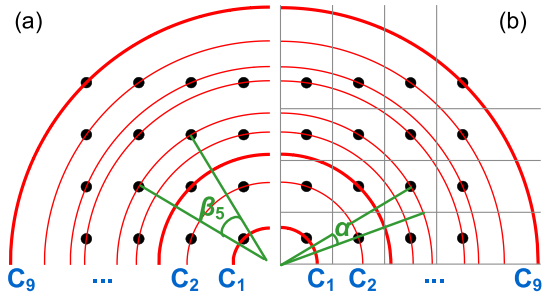

Fig. 2. Constellation points in a quadrant of 64-QAM: (a) definition of angular distance; (b) definition of phase margin.

TABLE I
ANGULAR DISTANCES AND PHASE MARGINS OF 64-QAM

Subset	Symbols	β_i (rad)	Phase margin (rad)
C_1	$\pm1 \pm 1i$	$\pi/2$	$\pi/4$
C_2	$\pm1 \pm 3i, \pm3 \pm 1i$	0.6435	0.3218
C_3	$\pm3 \pm 3i$	$\pi/2$	0.2945
C_4	$\pm1 \pm 5i, \pm5 \pm 1i$	0.3948	0.1974
C_5	$\pm3 \pm 5i, \pm5 \pm 3i$	0.4900	0.1903
C_6	$\pm5 \pm 5i$	0.2838	0.1841
	$\pm1 \pm 7i, \pm7 \pm 1i$		0.1419
C_7	$\pm3 \pm 7i, \pm7 \pm 3i$	0.7610	0.1392
C_8	$\pm5 \pm 7i, \pm7 \pm 5i$	0.3303	0.1366
C_9	$\pm7 \pm 7i$	$\pi/2$	0.1343

In the previous example, $3 + 5i$ are used to illustrate the ambiguity issue. However, not all the points are susceptible to this problem. As shown in Fig. 1(b), there will be only one minimizer for Eq. (4) when $1 + 1i$ are transmitted, under the condition that AWGN is small enough. To further investigate this issue, we take 64-QAM as an example and classify the constellation points into 9 subsets C_1, C_2, …, C_9 based on their magnitudes (see Fig. 2). For each subset, we define the minimum angular distance β_i as:

$$\beta_i = \min_{a_u, a_v \in C_i} |\arg(a_u) - \arg(a_v)| \tag{6}$$

It is worth mentioning that the angular distance discussed here is different from the concept of phase margin, which is defined as the minimum angle to rotate an ideal constellation point to the decision boundary of its adjacent points [17]. Fig. 2 (a) and (b) illustrates these two concepts, and the values of β_i and phase margins are summarized in Table I. Phase margins decrease monotonically as magnitude becomes larger, while that's not the case for β_i. For 64-QAM, only three subsets, C_1, C_3 and C_9 have β_i equal to $\pi/2$ and all other subsets have smaller angular distances. The worst case happens for C_6, where there are 12 points in this subset, making β_6 to be 0.2838 (rad). As the order of QAM increases, the angular distances will further shrink. For example, for 256-QAM, the minimum β_i is 0.1426 (rad).

B. V&V Algorithm

V&V algorithm, also known as *M*-th power, is originally proposed for *M*-ary PSK signals [16]. It exploits the rotational symmetry of the *M*-ary PSK constellation and raises the received signals to the *M*-th power to remove the data modulation. It cannot be directly applied to QAM formats, since such symmetry is lost. Lots of adaptations have been proposed, including single-stage and multi-stage algorithms [18]–[22]. Most of them are based on QPSK partitioning, where symbols with modulation angles of $\pi/4 + m \cdot \pi/2, m = 0, 1, \ldots$ are assigned more credits [18], [19] or used for coarse phase noise estimation in the first-stage [20]–[22]. It is because these symbols form separate QPSK constellation sets with different magnitudes and thus the data modulation can be eliminated by the fourth power operation. On the other hand, these symbols also have angular distances of $\pi/2$. Therefore, if the occurrences of these symbols are increased, the accuracy of CPE can be improved, no matter whether BPS or V&V based algorithm is adopted.

C. PS-QAM With Adaptable SNR and Laser Phase Noise Tolerance

Denote the symbol set as X, then the commonly used MB distribution generates a constellation point x with the probability [5]:

$$P_X(x) = \frac{e^{-\lambda|x|^2}}{\sum_{x' \in X} e^{-\lambda|x'|^2}} \tag{7}$$

where $\lambda \geq 0$ is the shaping factor that can be used to control the entropy $\boldsymbol{H}$ of the distribution. It is clear from Eq. (7) that in the MB distribution, points with smaller magnitudes are assigned higher probabilities. Although such assignment can achieve higher AIR in an SNR limited system, it doesn't take the relationship between the angular distance and the magnitude of the constellation point into consideration, as discussed in the previous subsection. For example, in the case of 64-QAM, points $\pm7 \pm 7i$ have angular distances equal to $\pi/2$ and should be generated with higher probabilities to combat laser phase noise, while they are allocated the lowest probabilities according to Eq. (7).

By comparison, MB/ADD distribution has the form:

$$P_X(x) = \frac{1}{Z} e^{-\lambda|x|^2} \cdot e^{\alpha\beta(x)} \tag{8}$$

where $\beta(x)$ is the function taking the angular distances of points x and α is another shaping factor to tune the entropy. Z is the normalization factor to ensure that the probabilities sum to 1 and is equal to $Z = \sum_{x' \in X} e^{-\lambda|x|^2} e^{\alpha \cdot \beta(x)}$. It is worth noting that the ADD part can take forms of other functions and here we choose the exponential function as a proof of concept.

Fig. 3 illustrates the probability mass functions (PMFs) of MB/ADD distributions under different λ and α. When $\alpha = 0$, as shown in Fig. 3(a), it becomes the conventional MB distribution. With larger α, MB/ADD distribution gets influenced by the ADD part more. It can also be viewed from Fig. 3(a), (b) and (c). The points $\pm3 \pm 3i$, which are supposed to have lower probabilities than those of $\pm3 \pm 1i$ in the MB distribution, gradually achieves higher probabilities, as α increases from 0 to 1.4. Fig. 3(d) depicts the case where $\lambda = 0$ and $\alpha = 1.4$, and it becomes the pure ADD distribution. By multiplying two families of distributions, MB/ADD distribution is equipped with two shaping factors, λ and α, to control the entropy. Fig. 4(a) shows the entropy as λ and α varies, where it can be clearly seen that there are multiple combinations of λ and α to get certain entropy. Therefore, a trade-off between better SNR tolerance

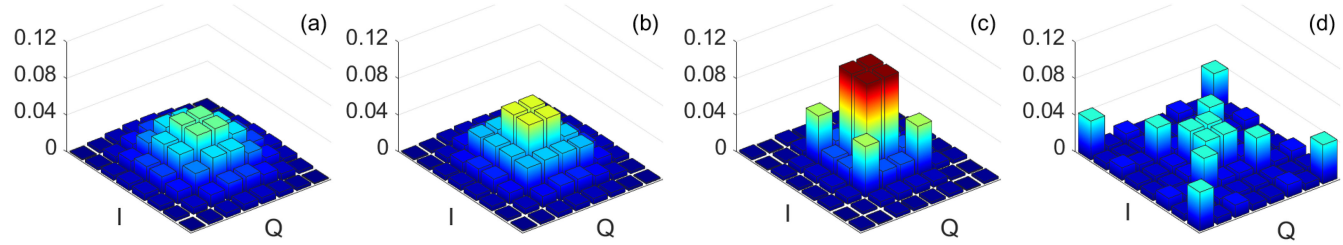

Fig. 3. PMFs of MB/ADD distributions: (a) $\lambda = 0.04$, $\alpha = 0$, $H = 5.47$ bits; (b) $\lambda = 0.04$, $\alpha = 0.03$, $H = 5.37$ bits; (c) $\lambda = 0.04$, $\alpha = 1.4$, $H = 4.69$ bits; (d) $\lambda = 0$, $\alpha = 1.4$, $H = 5.60$ bits.

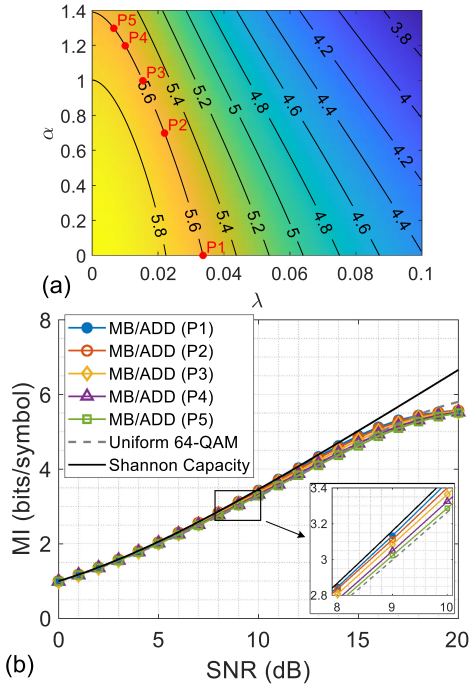

Fig. 4. (a) Entropy of MB/ADD distributions versus λ and α; (b) MI versus SNR for several MB/ADD distributions with $H = 5.6$ bits.

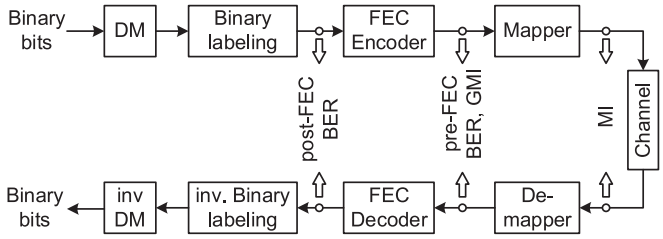

Fig. 5. Scheme of a communication system employing PAS.

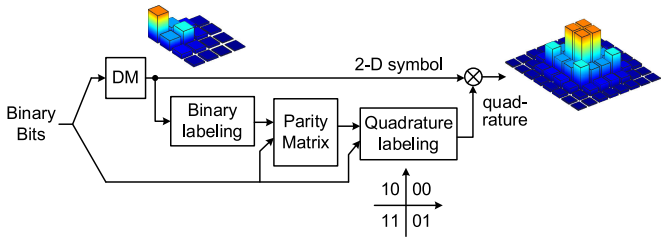

Fig. 6. Working principles of the 2-D PAS to generate PS-QAM with MB/ADD distribution.

and better laser phase noise tolerance can be made. In a system with severe AWGN and low laser phase noise, a relatively large λ and small α should be adopted. While in a system with small AWGN and severe laser phase noise, a MB/ADD distribution with relatively small λ and large α should be chosen. To further illustrate this idea, Fig. 4(b) depicts the mutual information (MI) of MB/ADD shaped 64-QAM in a memoryless AWGN channel. The MI is estimated from Monte Carlo simulations. All the MB/ADD distributions (P1–P5) have entropies of 5.6 bits and they are labelled on the Fig. 4(a) with red dots. As α increases, the gap to Shannon capacity increases and the shaping gain compared with uniform distribution decreases. However, as to be Illustrated in the Section III and V, this part of sacrificed shaping gain in the AWGN channel can be traded for improved laser phase noise tolerance.

D. Implementation

Fig. 5 shows the scheme of a communication system employing probabilistic amplitude shaping (PAS). The PAS adopts a reverse concatenation architecture, where a distribution matcher (DM) and an inverse DM (inv DM) are inserted before the FEC encoder and after the FEC decoder, respectively [23]. Commonly used DM algorithm includes constant composition distribution matching (CCDM) [24], hierarchical distribution matching (HiDM) [25], prefix-free code distribution matching (PCDM) [25], etc. To generate PS-QAM with MB distribution, the I and Q components are first generated independently using the PAS scheme and then paired to form a complex QAM symbol. For MB/ADD distribution, the I and Q components are not independent and thus this method cannot be directly applied. However, same as MB distribution, MB/ADD distribution also has reflective symmetry about the real and imaginary axes. Therefore, we can modify the original PAS scheme to generate 2-D symbols directly. Fig. 6 depicts the working principles, assuming that the FEC encoder adopts low density parity-check (LDPC) codes [23]. Similar to the 1-D version, the constellations are shaped on a per quadrature basis. Part of the uniformly distributed data bits and all of the FEC parity bits are used to determine the quadrature of the points. There are mainly two differences compared with the 1-D PAS. Firstly, every binary bits sequence is matched to a sequence of 2-D complex symbols, say $1 + 1i,\ 1 + 3i \ldots$, rather than 1, 3, …. Secondly, the quadrature of the 2-D symbol is determined directly by the quadrature label,

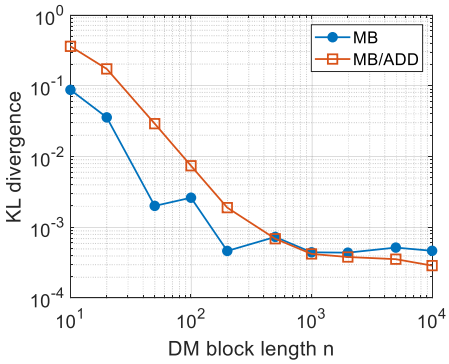

Fig. 7. KL divergence of MB and MB/ADD shaping when 1-D PAS and 2-D PAS are applied respectively.

rather than the Cartesian product of two signed labels. All the relationships between entropy, code rate and information rate derived in the original 1-D PAS can be directly applied to the 2-D version. For example, the code rate of the FEC encoder R_c for M-ary QAM in the 1-D PAS has to satisfy the relationship (suppose M is an even power of 2):

$$R_c \geq \frac{\log_2\sqrt{M}-1}{\log_2\sqrt{M}} \tag{9}$$

While in the 2-D case, it becomes:

$$R_c \geq \frac{\log_2 M-2}{\log_2 M} \tag{10}$$

which is the same as Eq. (9).

Compared with the conventional communication system employing uniform distribution, the increased complexity in the PAS scheme is mainly attributed to the DM and inv DM, which depends on the DM algorithm and the DM block length. A longer block leads to a more accurate implementation of the desired distribution and thus an increased shaping gain, yet at the cost of higher complexity. For CCDM, the complexity grows with $O(n)$ [26]. A practical solution to generate a long sequence of shaped QAM symbols is to concatenate the output of PAS of shorter block length. In the Fig. 7, we investigate this scenario by generating 100,000 shaped 64-QAM symbols when the block length n changes from 10 to 10,000. The target entropy for the DM is 5 bits. 1-D PAS and 2-D PAS are used for MB and MB/ADD distribution, respectively. The Kullback-Leibler (KL) divergence between the generated and the desired 2-D QAM distribution is calculated, where CCDM is assumed to be the DM algorithm. For a short block, e.g. $n < 100$, the 2-D PAS has larger KL divergence as it matches to a distribution of 16 2-D constellation points, while 1-D PAS matches to a distribution of 4 1-D amplitudes. This penalty of 2-D PAS vanishes as n increases. When block length is larger than 1000, these two distributions present similar performance. In other words, for KL divergence lower than 10^{-3}, the proposed MB/ADD shaping can be applied with the similar complexity when comparing with the conventional MB shaping. And as stated by [26], a block length of about 1000 symbols is quite feasible with fair implementation complexity and power consumption for application-specific integrated circuit (ASIC) implementation.

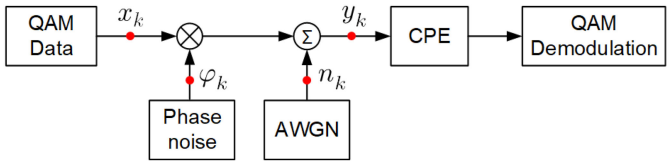

Fig. 8. Simulation setup.

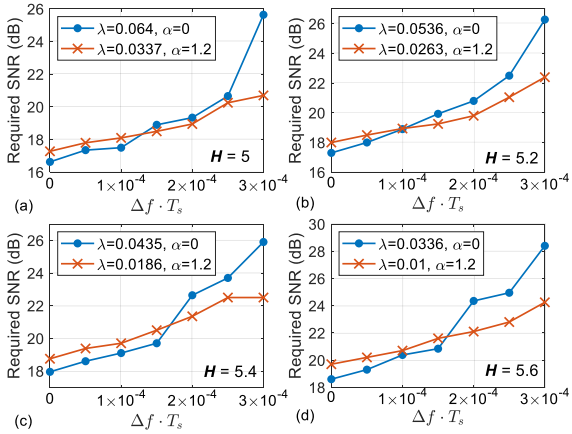

Fig. 9. Required SNR at pre-FEC BER of 4.5×10^{-3} versus laser phase noise for MB/ADD shaped 64-QAM under two parameter settings and under different entropies: (a) $H = 5$ bits; (b) $H = 5.2$ bits; (c) $H = 5.4$ bits; (d) $H = 5.6$ bits.

III. Simulation and Results

Firstly, the performance of the proposed MB/ADD distribution is investigated by simulation. Fig. 8 shows the simulation setup. Both phase noise and AWGN are considered. BPS is used in the CPE module. The number of test phases is 64 and block lengths are optimized respectively. Fig. 9(a)–(d) plots the required SNR at the 7% hard-decision (HD) pre-FEC BER threshold of 4.5×10^{-3} [28] for PS-64QAM adopting MB/ADD distribution (PS-64QAM MB/ADD) versus the $\Delta f \cdot T_s$ when $\boldsymbol{H}$ takes values of 5, 5.2, 5.4, 5.6. The pre-FEC BER is defined in Fig. 5 and the SNR is defined as the power of signals over the power of the loaded AWGN. Gray labeling is used in BER calculation. For each entropy, two sets of parameters are tested: i) $\alpha = 0$; ii) $\alpha = 1.2$. The exact values of λ and α are given in Fig. 9. Setting i) is a special case, where it becomes the conventional MB distribution. While in the setting ii), both MB and ADD parts are activated. When $\Delta f \cdot T_s < 1 \times 10^{-4}$, the major impairment is AWGN. Thus, setting i) has better performance. As $\Delta f \cdot T_s$ increases above 2×10^{-4}, the introduction of ADD part in the setting ii) is justified by achieving the same BER with lower SNR. Thus, setting ii), which is deficient in the MB shaping, enables more efficient use of the resources in the case of large phase noise. For example, when $\boldsymbol{H} = 5.2$ bits and $\Delta f \cdot T_s = 3 \times 10^{-4}$, SNR at the FEC threshold is reduced by 3.88 dB in the setting ii). Another example is to consider a case where $\Delta f \cdot T_s$ is 2.5×10^{-4}, SNR is 22.8 dB and 7% HD-FEC is adopted. By using setting ii), the entropy can be increased to 5.6 bits, while it is only 5.2 bits for setting i).

IV. Experimental Setup

We conduct experiments to further verify the performance of MB/ADD distribution. The experimental setup is shown in

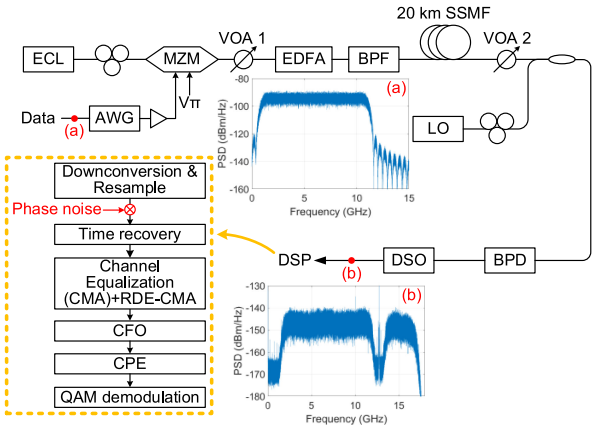

Fig. 10. Experimental setup and the signal spectra.

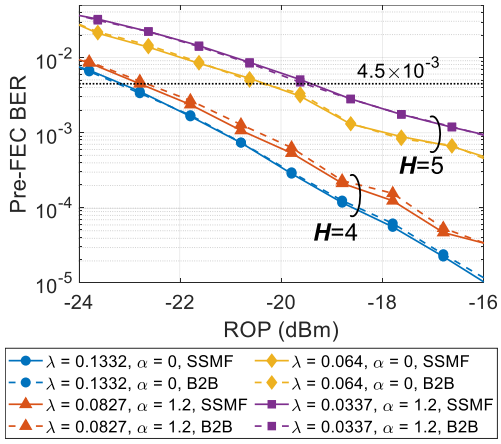

Fig. 11. Pre-FEC BER versus ROP for PS-64QAM MB/ADD (Solid line: after 20-km SSMF, dashed line: B2B).

Fig. 10. The transmitted symbols with symbol rate of 10 GBaud are first generated and pulse shaped by a root-raised-cosine filter (RRC) with a roll-off factor of 0.1. Then it is upconverted to an intermediate frequency (IF) of 6 GHz to make it a real signal, whose spectrum in shown in Fig. 10(a). An arbitrary waveform generator (AWG, Keysight M8195A) working at 32 GSa/s converts the digital samples into an analog waveform, which is modulated onto the optical carrier by a Mach-Zehnder modulator (MZM) biased at the null point. The optical carrier is kept unchanged throughout the experiments and it is from an external cavity laser (ECL, Pure Photonics PPCL100) with an intrinsic linewidth of 10 kHz. The modulated light is boosted by an erbium-doped fiber amplifier (EDFA) before entering a 20-km standard single mode fiber (SSMF). The variable optical attenuator 1 (VOA 1) is used to make sure that signals shaped by MB/ADD distribution with different parameter settings have the same power before being amplified by the EDFA. A 0.4-nm optical band-pass filter (BPF) is used to filter out out-of-band amplified spontaneous emission (ASE) noise. A LO mixes with the received signals via a 3-dB coupler and are then detected by a balanced photodetector (BPD, u2t BPDV2020R). A digital sampling oscilloscope (DSO, Keysight DSOZ254A) captures the heterodyne detected signals at a sampling rate of 40 GSa/s, the spectrum of which is shown in Fig. 10(b). The offline DSP is implemented in the Matlab, where down-conversion and down-sampling, time recovery, channel equalization based on radius directed constant modulus algorithm (RDE-CMA), carrier frequency offset (CFO), CPE and QAM demodulation are performed sequentially. Under low received optical power (ROP), CMA is also included for pre-convergence.

When the odd-order QAM or PS-QAM symbols based on MB distribution are transmitted, utilizing the traditional Fast Fourier transform (FFT) based CFO may fail to get an accurate estimation of the frequency offset (FO). Usually, the FO is found by locating the maximum magnitude peak of the spectrum when the received signals are raised to the fourth power. It exploits the fourth-order non-circular property of square QAM formats, i.e. $\boldsymbol{E}_{\boldsymbol{X}}[x_k^4] \neq 0$ [29]. However, in odd-QAM or PS-QAM symbols, the $\boldsymbol{E}_{\boldsymbol{X}}[x_k^4]$ suffers from degradation as a result of the lack [30] or lower probabilities of symbols with large magnitudes [31]. The magnitude of the peak in the spectrum will further decrease when there is large laser phase noise in the system. Here, to solely evaluate the impact of different distributions on the CPE, we apply the modified FFT based CFO proposed in [30], where symbols located on the QPSK rings are selected and amplified before being raised to the fourth power. For 32-QAM symbols, the points $\pm 1 \pm 1i$ and $\pm 3 \pm 3i$ are amplified, while $\pm 1 \pm 1i$, $\pm 3 \pm 3i$ and $\pm 7 \pm 7i$ are amplified in the case of 64-QAM.

V. Results and Discussion

Our experiments consist of two parts, where we evaluate the SNR tolerance and laser phase noise tolerance of MB/ADD distribution, respectively.

A. SNR Tolerance

The SNR tolerance is tested when the system impairments are dominated by AWGN and the LO is another an ECL (Pure Photonics PPCL100) with an intrinsic linewidth of 10 kHz. Fig. 11 shows the pre-FEC BER of PS-64QAM MB/ADD. Here we experiment with two cases: $\boldsymbol{H} = 4$ bits and $\boldsymbol{H} = 5$ bits. Same as before, there are two sets of parameters, i) $\alpha = 0$ and ii) $\alpha = 1.2$. Both settings exhibit better performance with higher ROP, as a result of increased SNR. Compared with back-to-back (B2B) scenario, 20-km SSMF transmission introduces minimal power penalty. In this experiment, all the other impairments play insignificant role and thus the system operates under the SNR-limited condition, which leads to lower BER of setting i). At the pre-FEC BER threshold of 4.5×10^{-3}, power penalty of setting ii) compared with setting i) is about 0.1 dB and 0.8 dB for $\boldsymbol{H} = 4$ bits and $\boldsymbol{H} = 5$ bits, respectively.

B. Laser Phase Noise Tolerance

The laser phase noise tolerance is evaluated by two steps. In the first step, we investigate the performance of setting i) and ii) as laser phaser noise varies. This is achieved by using one narrow linewidth ECL as the LO, which is the same as that in the previous subsection, and digitally adding phase noise to the received signals. Phase noise are generated offline according to Eq. (2), with Δf changing to simulate different phase noise. They are added to the signals in the DSP module, right after down-conversion, as shown in Fig. 10. In the second step,

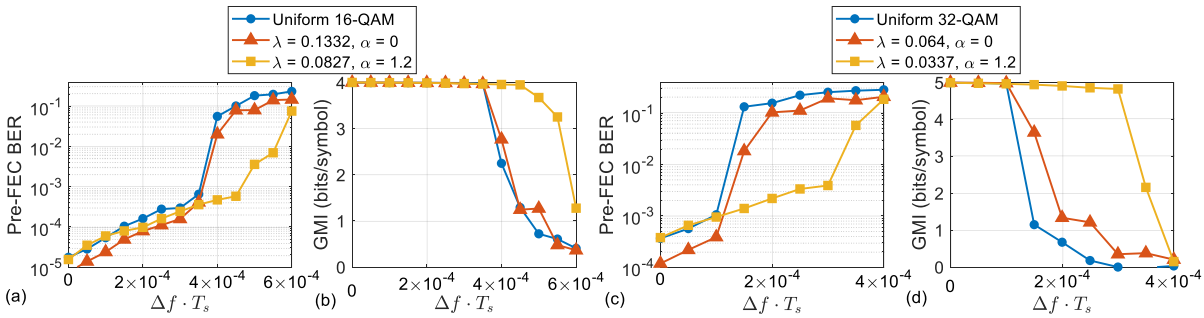

Fig. 12. Pre-FEC BER and GMI for uniform 16-QAM, uniform 32-QAM and PS-64QAM MB/ADD under two settings versus digitally added phase noise: (a)–(b): $H = 4$ bits; (c)–(d): $H = 5$ bits.

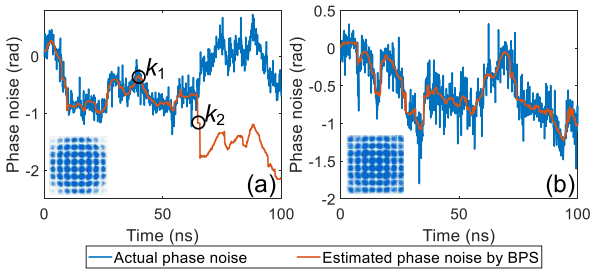

Fig. 13. Actual and estimated phase noise with insets showing constellation diagrams: (a) setting i); (b) setting ii).

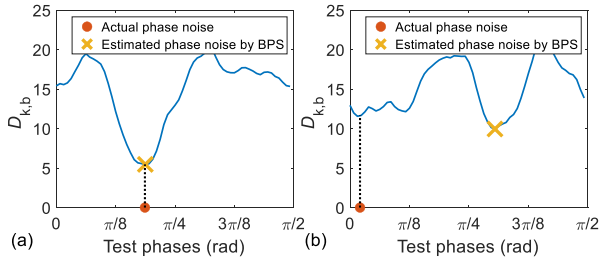

Fig. 14. $\boldsymbol{D_{k,b}}$ versus actual phase noise, at instant (a) k1; (b) k2.

to further verify the improved laser phase noise tolerance of MB/ADD distribution, we replace the LO with a DFB laser (Narda-MITEQ SCMT-100M13G), whose linewidth is several mega-hertz. All measurements are executed after 20-km SSMF transmission.

1) Digitally Generated Phase Noise: Fig. 12 shows the pre-FEC BER and GMI of PS-64QAM MB/ADD versus $\Delta f \cdot T_s$ when $\boldsymbol{H} = 4$ bits and $\boldsymbol{H} = 5$ bits. Same shaping parameters from the previous subsection is used. BPS is used for CPE at this stage. Uniform 16-QAM and 32-QAM are also shown for references. Similar with simulation results, setting ii) presents better phase noise tolerance with the increase of $\Delta f \cdot T_s$, which is manifested by lower pre-FEC BER and higher GMI in Fig. 12(a)–(d). The improvement is more notable in the case of $\boldsymbol{H} = 5$ bits. When $\Delta f \cdot T_s$ is larger than 1.5×10^{-4}, signals of setting i) can be barely demodulated, while setting ii) can still achieve pre-FEC BER of 1.39×10^{-3}. At the pre-FEC BER threshold of 4.5×10^{-3}, $\Delta f \cdot T_s$ for setting i) is about 3.8×10^{-4} and 1.3×10^{-4} for $\boldsymbol{H} = 4$ bits and $\boldsymbol{H} = 5$ bits, respectively. While they are about 5.2×10^{-4} and 3.0×10^{-4} in the setting ii). As a result, laser linewidth requirements can be significantly relaxed by choosing parameters properly in the MB/ADD based shaping. For example, for 40 GBaud PS-64QAM with $\boldsymbol{H} = 5$ bits, a 12-MHz linewidth LO can be employed in the setting ii) under the assumption of 7% HD-FEC.

2) DFB Generated Phase Noise: Fig. 13 illustrates the actual phase noise over time, acquired by comparing the received signals before CPE with the transmitted signals, and the phase noise estimated by BPS. Constellation diagrams are shown in the insets. ROP is fixed at -12.9 dBm and $\boldsymbol{H}$ is set to 5 bits. Still, we are comparing setting i) and ii). In the setting i), the BPS fails to track the rapid phase changes under this condition, as shown by the discrepancy of actual and estimated phase noise. By comparison, setting ii) still enables an accurate phase noise estimation, although the LO has large linewidth. Consequently, pre-FEC BER for setting i) is 1.17×10^{-1}, while it is 3.84×10^{-3} for setting ii).

To better understand the misestimation of BPS in setting i), $D_{k,b}$ at instants k_1 and k_2 (see Fig. 13(a)) are calculated and plotted in the Fig. 14. As explained in the Section II, the phase noise is estimated by finding the global minimum of $D_{k,b}$,which is marked as the " × " in the graph. The actual phase noise is marked as ".". Compared with instant k_1, $D_{k_2,b}$ has several local minima, which makes it quite challenging to accurately find the phase noise.

Pilot symbols can be used to assist in tracking phase noise. One pilot symbol is inserted periodically after $P-1$ data symbols, providing a pilot rate of $1/P$. Here, we adopt the two-stage CPE, the same as [32]. In the first stage, the received signals are multiplied by a zero-padded pilot sequence p_k, which is defined as:

$$p_k = \begin{cases} x_k, k = iP, \ i = 1, 2, 3 \ldots \\ 0, \qquad\qquad \text{otherwise} \end{cases} \tag{11}$$

The resulting sequence is sent into an interpolation filter, where two moving average (MA) filters with the same tap number are cascaded. The first-stage CPE is acquired by taking the argument of the filtered sequence. For the second-stage CPE, a traditional BPS is applied. In our experiment, P takes on values of 50, 25, 20, 16, 12, 10, making pilot rates of 2%, 4%, 5%, 6.25%, 8.33% and 10%. Fig. 15 plots the pre-FEC BER versus pilot rate for setting i). When pilot rates are lower than 4%, although setting i) can benefit from an improved accuracy of CPE and BER decreases accordingly, it is still worse than that of setting ii) with no pilot symbols inserted. The benefits of pilot symbols can only be obtained with pilot rates higher than 5%.

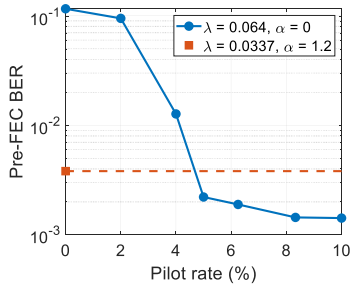

Fig. 15. Pre-FEC BER of PS-64QAM MB/ADD with setting i) under different pilot rates (setting ii) without pilot symbols shown as reference).

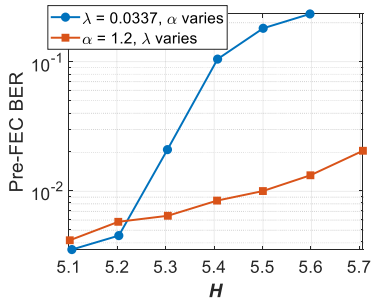

Fig. 17. Pre-FEC BER for PS-64QAM MB/ADD as H varies.

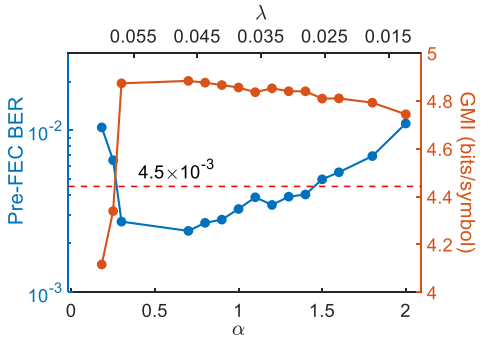

Fig. 16. Pre-FEC BER and GMI versus $\boldsymbol{\lambda}$ and $\boldsymbol{\alpha}$ with $H = 5$ bits.

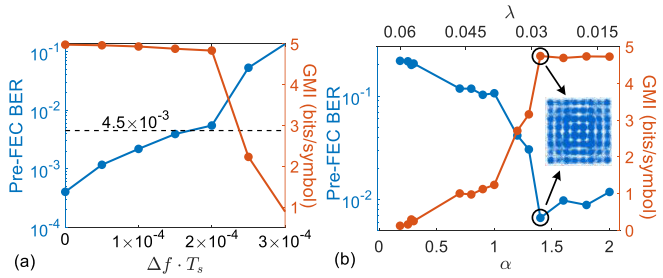

Fig. 18. Pre-FEC BER and GMI of two-stage V&V based CPE: (a) versus $\boldsymbol{\Delta f \cdot T_s}$; (b) versus $\boldsymbol{\lambda}$ and $\boldsymbol{\alpha}$ with $H = 5$ bits (insets showing constellation diagram when $\boldsymbol{\lambda} = 0.0179$, $\boldsymbol{\alpha} = 1.3$).

Thus, setting ii) presents better laser phase noise tolerance when pilot rate is constrained to 4%.

As shown in Section III, increasing SNR can also lead to lower BER, as the impact of AWGN on the CPE can be reduced. When ROP is increased by 3.9 dBm to −9.0 dBm, pre-FEC BER for setting i) is decreased to 5.56×10^{-2}. It is still worse than that of setting ii) when ROP is −12.9 dBm. What's more, adopting higher ROP can also improve the performance of setting ii), as BER is reduced to 2.96×10^{-3}. Thus, increasing SNR can only bring about limited improvements to setting i) when the coherent receiver suffers from large phase noise.

In the previous discussion, we only explore the performance of MB/ADD shaping with two settings, i.e. α takes values of 0 and 1.2, respectively. In the following experiment, we measure the performance of PS-64QAM MB/ADD under different λ and α, fixing $\boldsymbol{H}$ at 5 bits. The results are shown in Fig. 16. When $\alpha < 0.7$, increasing α leads to lower pre-FEC BER and higher GMI, due to the better performance of ADD distribution in the scenario of large laser phase noise. However, BER increases as α increases beyond 0.7. This is because the system is now under the combined impact of AWGN and laser phase noise. Further increase in α give rises to better phase noise tolerance, yet at the expense of sacrificing SNR tolerance, as MB distribution can make better use of the limited signal power to approach Shannon capacity. If the system is targeted at certain entropy, Fig. 16 can be used as a guidance on choosing proper values of λ and α. A trade-off between SNR tolerance and laser phase noise tolerance can be achieved, allowing the system to agilely adapt to various channel conditions, including launched power and laser linewidth.

Similarly, both λ and α can be changed to attain higher entropy. For the measurements shown in Fig. 11, only λ is varied and α is kept at 1.2 when $\boldsymbol{H}$ is increased to 5 bits in the setting ii). As a result, the power penalty of setting ii) compared with setting i) is higher in the case of $\boldsymbol{H} = 5$ bits than that of $\boldsymbol{H} = 4$ bits. It indicates that when the laser phase noise exerts minimal impact on the received signals, changing α is a better choice to increase SE. However, things are different when a DFB laser is used as the LO. Fig. 17 investigates this scenario and compares pre-FEC BER versus $\boldsymbol{H}$ when only one of the λ and α is varied, keeping λ at 0.0337 or α at 1.2, respectively. Unlike the case of varying α, varying λ produces a BER-$\boldsymbol{H}$ curve with smaller slope. Therefore, under large laser phase noise, modifying MB/ADD distribution by changing λ enables a smoother evolution towards higher SE.

3) V&V Based CPE: All the discussions above are based on BPS, which achieves better performance than V&V based algorithm at the cost of higher complexity. However, as mentioned in Section II, increasing the probabilities of points with higher angular distances would also be beneficial to QPSK partitioning based V&V algorithm. In this step, we revisit the data captured in the previous two steps, and replace the BPS with two-stage V&V based CPE [20]. In the first stage, the received signals are partitioned and those belong to subsets C_1, C_3 and C_9 are selected to be processed by a traditional V&V algorithm to get a coarse CPE. For signals belong to other subsets, their estimated phase noise are interpolated. In the second stage, a Maximum-Likelihood (ML) method [33] is conducted to obtain a fine CPE.

We first evaluate the two-stage V&V based CPE under digitally added phase noise. Fig. 18(a) shows the pre-FEC BER and GMI of PS-64QAM MB/ADD with setting ii) and $\boldsymbol{H} = 5$ bits, as $\Delta f \cdot T_s$ varies. Even if a simplified CPE is employed, it still presents better laser phase noise tolerance than setting i) with BPS applied. At $\Delta f \cdot T_s$ of 1.5×10^{-4}, pre-FEC BER is 3.93×10^{-3}. And the $\Delta f \cdot T_s$ at BER threshold of 4.5×10^{-3} is about 1.7×10^{-4}.

Then we move to the case of DFB generated phase noise. Fig. 18(b) depicts pre-FEC BER and GMI of PS-64QAM MB/ADD versus λ and α when $\boldsymbol{H}$ remains 5 bits. In comparing with Fig. 16, it presents different behaviors than that of BPS. Pre-FEC BER and GMI become better as α increases. When α is larger than 1.8, it achieves similar performance with BPS. This is because the probabilities of points in subsets C_1, C_3 and C_9 become higher with larger α. Consequently, there are more symbols involved in the first stage CPE and thus a more accurate estimation can be made. Although pre-FEC BER is worse than that of BPS, by choosing appropriate λ and α, it is still possible to achieve error-free transmission by using soft decision FEC (SD-FEC) with higher overhead.

VI. Conclusion

In this paper, a novel MB/ADD distribution is proposed for PS-QAM in coherent communication system. By overlaying the ADD distribution on top of the traditional MB distribution, the laser phase noise and SNR adaptability can be achieved concurrently. The performance of MB/ADD distribution is verified through simulations and experiments, where different sets of parameters are tested. Results show that apart from preserving the improvement of MB distribution in the SNR limited region, MB/ADD distribution also assures the benefits of PS in the large phase noise region. In the simulation analysis, MB/ADD shaping present lower SNR requirements when evaluated at the same pre-FEC BER threshold with MB shaping, under severe phase noise. In the experiment analysis, MB/ADD shaping can reduce pre-FEC BER by one order of magnitude when large linewidth lasers are employed, without sacrificing data rate for pilot symbols. And the improvement can be observed in both BPS based and V&V based CPE. What's more, MB/ADD shaping can be implemented with the similar complexity with MB shaping, if the KL divergence between the generated and the expected 2-D QAM distribution is required to be lower than 10^{-3}. Therefore, the proposed PS-QAM with MB/ADD distribution is a promising candidate for flexible low-cost coherent communication systems.

References

[1] K. Wang *et al.*, "Dual-carrier 400G field trial submarine transmission over 6,577-km using 60-GBaud digital faster-than-Nyquist shaping PDM-QPSK modulation format," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2015, Paper W3E.2.

[2] K. Schuh *et al.*, "Single carrier 1.2 Tbit/s transmission over 300 km with PM-64 QAM at 100 GBaud," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2017, Paper Th5B.5.

[3] G. Bosco, "Advanced modulation techniques for flexible optical transceivers: The rate/reach tradeoff," *J. Lightw. Technol.*, vol. 37, no. 1, pp. 36–49, Jan. 2019.

[4] F. Buchali *et al.*, "Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration," *J. Lightw. Technol.*, vol. 34, no. 7, pp. 1599–1609, Apr. 2016.

[5] J. Cho and P. J. Winzer, "Probabilistic constellation shaping for optical fiber communications," *J. Lightw. Technol.*, vol. 37, no. 6, pp. 1590–1607, Mar. 2019.

[6] I. B. Djordjevic *et al.*, "Coded polarization multiplexed iterative polar modulation (PM-IPM) for beyond 400 Gb/s serial optical transmission," in *Proc. Opt. Fiber. Conf.*, Mar. 2010, Paper OMK2.

[7] T. H. Lotz *et al.*, "Coded PDM-OFDM transmission with shaped 256-iterative-polar-modulation achieving 11.15-b/s/Hz intrachannel spectral efficiency and 800-km reach," *J. Lightw. Tehcnol.*, vol. 31, no. 4, pp. 538–545, Feb. 2013.

[8] R. Krishnan *et al.*, "Constellation optimization in the presence of strong phase noise," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 5056–5066, Dec. 2013.

[9] T. Fehenberger *et al.*, "Sensitivity gains by mismatched probabilistic shaping for optical communication systems," *IEEE Photon. Technol. Lett.*, vol. 28, no. 7, pp. 786–789, Apr. 2016.

[10] F.R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 913–929, May 1993.

[11] S. L.I. Olsson *et al.*, "Record-high 17.3-bit/s/Hz spectral efficiency transmission over 50 km using probabilistically shaped PDM 4096-QAM," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2018, Paper Th4C.5.

[12] T. Sasai *et al.*, "Laser phase noise tolerance of uniform and probabilistically shaped QAM signals for high spectral efficiency systems," *J. Lightw. Tehcnol.*, vol. 38, no. 2, pp. 439–446, Jan. 2020.

[13] F. Pittalà *et al.*, "400-Gbit/s DP-16-QAM transmission over 40-km unamplified SSMF with low-cost PON lasers," *IEEE Photon. Technol. Lett.*, vol. 31, no. 15, pp. 1229–1232, Aug. 2019.

[14] E. Ip and J. M. Kahn, "Feedforward carrier recovery for coherent optical communications," *J. Lightw. Tehcnol.*, vol. 25, no. 9, pp. 2675–2692, Sep. 2007.

[15] T. Pfau, S. Hoffmann, and R. Noé. "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for M-QAM constellations," *J. Lightw. Tehcnol.*, vol. 27, no. 8, pp. 989–999, Apr. 2009.

[16] A. J. Viterbi and Audrey M. Viterbi. "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Trans. Inf. Theory*, vol. 29, no. 4, pp. 543–551, Jul. 1983.

[17] S. Okamoto *et al.*, "Laser phase noise tolerance of probabilistically-shaped constellations," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2018, Paper W2A.51.

[18] M. Seimetz, "Laser linewidth limitations for optical systems with high-order modulation employing feed forward digital carrier phase estimation," in *Proc. Opt. Fiber Commun. Conf.*, Feb. 2008, Paper OTuM2.

[19] I. Fatadin, D. Ives, and S. J. Savory, "Laser linewidth tolerance for 16-QAM coherent optical systems using QPSK partitioning," *IEEE Photon. Technol. Lett.*, vol. 22, no. 9, pp. 631–633, May 2010.

[20] Y. Gao *et al.*, "Low-complexity two-stage carrier phase estimation for 16-qam systems using qpsk partitioning and maximum likelihood detection," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2011, Paper OMJ6.

[21] S. M. Bilal *et al.*, "Multistage carrier phase estimation algorithms for phase noise mitigation in 64-quadrature amplitude modulation optical systems," *J. Lightw. Tehcnol.*, vol. 32, no. 17, pp. 2973–2980, Sep. 2014.

[22] S. M. Bilal *et al.*, "Carrier phase estimation through the rotation algorithm for 64-QAM optical systems," *J. Lightw. Tehcnol.*, vol. 33, no. 9, pp. 1766–1773, May 2015.

[23] G. Böcherer, F. Steiner, and P. Schulte, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4651–4665, Dec. 2015.

[24] P. Schulte and G. Böcherer, "Constant Composition Distribution Matching," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 430–434, Jan. 2016.

[25] T. Yoshida, M. Karlsson, and E. Agrell, "Hierarchical distribution matching for probabilistically shaped coded modulation," *J. Lightw. Technol.*, vol. 37, no. 6, pp. 1579–1589, Mar. 2019.

[26] J. Cho, "Prefix-free code distribution matching for probabilistic constellation shaping," *IEEE Trans. Commun.*, vol. 68, no. 2, pp. 670–682, Feb. 2020.

[27] H. Sun *et al.*, "800G DSP ASIC design using probabilistic shaping and digital sub-carrier multiplexing," *J. Lightw. Tehcnol.*, to be published, doi: 10.1109/JLT.2020.2996188.

[28] "OTU4 long-reach interface," *ITU-T G.709.2, ITU*, Geneva, Switzerland, Jul. 2018.

[29] B. Picinbono, "On circularity," *IEEE Trans. Signal Process*, vol. 42, no. 12, pp. 3473–3482, Dec. 1994.

[30] F. Xiao *et al.*, "Feed-forward frequency offset estimation for 32-QAM optical coherent detection," *Opt. Express*, vol. 25, no. 8, pp. 8828–8839, Apr. 2017.

[31] Q. Yan, L. Liu, and X. Hong, "Blind carrier frequency offset estimation in coherent optical communication systems with probabilistically shaped M-QAM," *J. Lightw. Tehcnol.*, vol. 37, no. 23, pp. 5856–5866, Dec. 2019.

[32] M. Magarini *et al.*, "Pilot-symbols-aided carrier-phase recovery for 100-G PM-QPSK digital coherent receivers," *IEEE Photon. Technol. Lett.*, vol. 24, no. 9, pp. 739–741, May 2012.

[33] X. Zhou, "An improved feed-forward carrier recovery algorithm for coherent receivers with M-QAM modulation format," *IEEE Photon. Technol. Lett.*, vol. 22, no. 14, pp. 1051–1053, Jul. 2010.