Rate Redundancy and Entropy Allocation for PAS-OFDM Based Mobile Fronthaul

Rui Zhang, You-Wei Chen, Wenlong Mou, and Gee-Kung Chang, *Fellow, IEEE*

***Abstract*—From the perspective of an experimentalist, we discuss the comparison framework that ensures the same information rate in probabilistic shaping (PS) signals. Theoretically, the information rate can be tuned by adjusting the target entropy, scaling the signal bandwidth and adjusting FEC code rate. In practice, the information rate tuning should be based on the hardware feasibility (e.g., DAC resolution, available analog bandwidth, cost of rate adaptation). Adjusting a constellation entropy rate without considering FEC redundancy will cause an over-estimation of the PS system performance gain. To avoid the over-estimation issue, the modified pre-FEC threshold for PS signal is also proposed to estimate the system performance accurately. The modified pre-FEC threshold is obtained from the FEC code rate of PS signal. Moreover, we experimentally investigated the adaptive entropy allocation scheme in a millimeter-wave analog radio over fiber fronthaul using the modified pre-FEC threshold. By grouping the subcarriers into PS units and using the same QAM order, the proposed scheme can reduce the required data frame length as well as the processing complexity. Notably, this method requires neither dynamic FEC coding rate nor bandwidth adjusting method. Up to 2.5-dB power margins and a smoother pre-FEC performance among PS units are achieved in a frequency selective fading channel. If the modified FEC threshold were not used, it would induce up to 0.7-dB over-estimation in the received sensitivity. This indicates the importance of employing the proposed methods for comparison when FEC is used in combination with PS.**

***Index Terms*—Forward error correction (FEC) redundancy, millimeter wave, mobile fronthaul, probabilistic shaping, radio-over-fiber.**

I. Introduction

RECENTLY there have been substantial research progress in probabilistic shaping (PS) in both data center and mobile access applications [1]–[11]. In a system under average power constraints, PS signal with the Maxwell-Boltzmann distribution increases the Euclidean distances between adjacent transmitted symbols, and thus is more tolerant to amplified spontaneous emission (ASE) noise compared with the uniform distribution [12]. On the other hand, in the millimeter-wave (MMW) based radio-over-fiber fronthaul, the orthogonal frequency-division multiplexing (OFDM) signal suffers from a non-uniform channel response due to frequency selective power fading or local oscillator leakage [13]. PS enables continuous tuning of the entropies in comparing to the bit loading, which yields better system performance [6], [14], [15]. Considering the above two performance metrics, PS is an attractive solution in the OFDM based mobile fronthaul. In practice, it is very costly to perform rate adaptation based on SNR [16] so a fixed FEC code rate is preferred. So far, some research works are conducted with a fixed constellation entropy rate, which is defined as the product of signal entropy and signal bandwidth, to exploit the performance gain from the PS signals [1]–[5], [7], [9]–[11], [17] while others are maximizing the achievable data rate by applying PS signals. In the former case, it is common to compare the performance by scaling the signal bandwidth with a constant factor, under a fixed constellation entropy rate. The constant factor is defined as the ratio between the entropy of uniform signal and that of the PS signal, where the entropy is in the unit of bits/symbol. However, as discussed in [18], [19], unlike the calculation in conventional serially concatenated systems using the uniform signal, the information rate cannot be obtained from simple multiplication in PS systems. In this paper, we show that in a system with a fixed constellation entropy rate, the FEC induced redundancy, which yields a lower information rate for PS signal, needs to be considered carefully.

In this paper, we discuss several settings for conducting a fair comparison of PS systems under different operation senarios. We also experimentally evaluate a novel entropy allocation scheme based on PS units [15] based on a constant composition distribution matcher (CCDM). By grouping several subcarriers into one PS unit for entropy loading, this scheme relaxes the required symbol length as compared to per subcarrier entropy mapping scheme, and thus it reduces latency in wireless applications [15]. Moreover, the proposed scheme, containing the same QAM level in each PS unit and a fixed target entropy, does not require dynamical adjustment on either the FEC code rate or the bandwidth, and it results in a reduction of the processing complexity. This scheme enables a flexible SNR tracking by adjusting the PS unit size (number of subcarriers per PS unit) compared with the case where a uniform PS setting is applied to all subcarriers [1], [17]. Note that when

Manuscript received January 4, 2020; revised March 14, 2020; accepted April 13, 2020. Date of publication April 20, 2020; date of current version July 28, 2020. This work was supported by an NSF grant Award 1821819 to Industry/University Cooperative Resrearch Center for Fiber Wireless Integration and Networking (FiWIN) at Georgia Institute of Technology *(Corresponding author: Rui Zhang.)*

Rui Zhang, You-Wei Chen, and Gee-Kung Chang are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30308 USA (e-mail: ruizhangece@gatech.edu; yu-wei.chen@ece.gatech.edu; gkchang@ece.gatech.edu).

Wenlong Mou is with the Department of Electrical Engineering and Computer Science, University of California at Berkeley, Berkeley, CA 94720 USA (e-mail: wmou@eecs.berkeley.edu).

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

Digital Object Identifier 10.1109/JLT.2020.2988163

the target distribution has a structure of product distribution, the product distribution matcher (PDM) algorithm [20] can be incorporated into the proposed PS unit scheme to further reduce the computational complexity. In this paper, we work with Maxwell-Boltzmann distributions and focus on CCDM algorithm for simplicity. We evaluate the received sensitivity of the entropy allocation in a millimeter-wave radio-over-fiber mobile fronthaul using 64/256QAM OFDM signals and the modified pre-FEC thresholds for conducting a fair comparison. Different levels of performance over-estimation in the fronthaul link, such as failing to maintain the same information rate, are experimental demonstrated.

This paper is organized as follows. Section II describes the principle of operation. It includes a review of state-of-art probabilistic amplitude shaping (PAS) scheme, FEC induced redundancy, net bit rate analysis with proposed solutions to maintain in different application scenarios, and entropy allocation with PS units. Section III presents the experimental setup of the MMW based fronthaul while Section IV demonstrates the experimental results with modified pre-FEC thresholds. Finally, Section V gives concluding remarks.

II. Principle of Operation

In this section, we review the principle of operation using the state-of-art PAS scheme of constant composition distribution matcher (CCDM) and the proposed entropy allocation based on the PS unit (grouped subcarriers) size. Section A revisits the definition of spectral efficiency, information rate, and constellation entropy rate in a PAS-OFDM system, and their relationships. Section B and C discuss and analyze the possible cases of over-estimation as well as hardware requirements in an experimental implementation.

Note that in an experimental investigation, it would be more convenient to compare the performance using pre-FEC metrics since post-FEC BER requires more computational complexity and many investigations are based on un-coded systems without any FEC. Pre-FEC BER or Q^2 factor is widely used in industry standard with a threshold to ensure the a certain post-FEC BER in an un-coded system. In Section B and C, different cases of FEC with different net coding gains (NCG) are investigated for proof of concept. The FECs are Staircase FEC (defined in ITU-T G.709.2), openFEC (oFEC, standardized by Open ROADM targeting metro applications), and Concatenated FEC [21] (cFEC, defined in OIF 400ZR standard). Table I summarizes the Pre-FEC Q^2 factor thresholds for Uniform signal with the post-FEC BER setting as 10^{-15}. We will utilize the Q_1, Q_2, Q_3 and Q_4 in both the following analysis and experiment for uniform QAM signals. Q_4 with a heavier FEC OH is demonstrated for experimental investigation of the trend with an increased OH. Normalized generalized mutual information (NGMI) is another pre-FEC metric which is used for soft-decision FEC. NGMI thresholds of 0.8 (25% OH for ideal FEC) and 0.8621 (16% FEC OH for ideal FEC) are also discussed. These pre-FEC thresholds for uniform QAM signals will also be utilized in the experimental analysis as discussed in Section IV.

TABLE I
Pre-FEC Q^2 Factor Thresholds for Uniform QAM Signal

	Threshold value	OH	FEC scheme
Q_1	6.2509 dB	15.3%	oFEC
Q_2	7.0466 dB	14.8%	cFEC
Q_3	8.3369 dB	6.69%	Staircase FEC
Q_4	5.758 dB	30%	oFEC

A. Spectral Efficiency, Information Rate and Constellation Entropy Rate of PAS-OFDM Signal

PAS is the state-of-art scheme to preserve the target distribution from the distribution matcher (DM) with FEC encoding. PAS works for the case that the 2-D distribution is symmetric along the I axis and Q axis, and the probabilities of I and Q are independent with each other.

Fig. 1 illustrates the block diagram of the PAS scheme. M-QAM symbols require two parallel PAS blocks to generate I part and Q part (two PAM signal), respectively. Firstly, a pseudorandom binary sequence (PRBS) with a length of $\mathrm{k}+\gamma n$ is divided into sequence 1 with a length of k and sequence 2 with a length of γn, respectively. Sequence 1 is loaded into the DM to generate the desired amplitude distribution. Given the targeted probabilities of the I part $(P_I(A))$ and Q part $(P_Q(A))$ in a PS M-QAM symbol, a target entropy of the QAM constellation can be expressed as $\mathbb{H}(X)=\mathbb{H}(P_I)+\mathbb{H}(P_Q)$. For each PAS block in Fig. 1, the target entropy of the DM is $\mathbb{H}(P_I)-1$ or $\mathbb{H}(P_Q)-1$ with one bit serving as sign part.

Most experimental demonstrations employ Maxwell-Boltzmann (MB) distribution to achieve enhanced Euclidean distance in an averaged power limited system. The probabilities are proportional to $\exp(-\lambda\cdot A^2)$, where *A* denotes the symbol amplitude. Thus, in this special case, I and Q follow the same distribution and are independent, which yields $\mathbb{H}(P_I)=\mathbb{H}(P_Q)=\frac{1}{2}\mathbb{H}(X)$.

In the following, assuming MB distribution, we calculate the spectral efficiency of the given system, and decompose the error between spectral efficiency and the target entropy into rate losses and overheads from different components in the system.

First, the output of DM serves as the amplitude of the PS-PAM symbol. When CCDM is used, a rate loss is induced by the DM scheme. In particular, the rate of the output of the distribution matcher is [19], [22]:

$$R_{dm}=(n_1+n_3+\cdots+n_{\sqrt{M}-1})^{-1}\log_2\frac{(n_1+n_3+\cdots+n_{\sqrt{M}-1})!}{n_1!n_3!\cdots n_{\sqrt{M}-1}!}, \quad (1)$$

where $n_i$is the number of symbols with each amplitude level in a PAS block. This DM rate is known to be strictly smaller

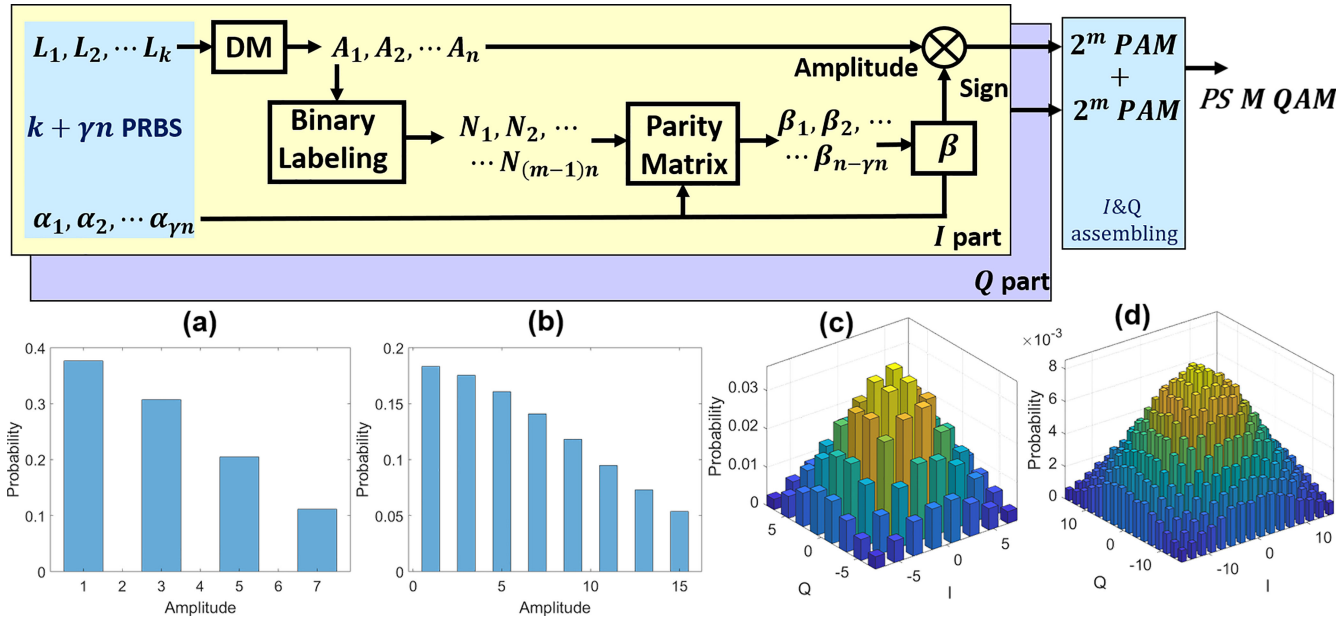

Fig. 1. Block diagram of the PAS scheme. PRBS: Pseudorandom binary sequence. Maxwell-Boltzmann distribution-based PAS: Probability distribution of the amplitude part with (a) $m = 3$, $H = 5.75$ bits/symbol, and (b) $m = 4$, $H = 7.8$ bits/symbol. 2-D probability distribution of the assembled QAM symbols with (c) $m = 3$, $H = 5.75$ bits/symbol, and (d) $m = 4$, $H = 7.8$ bits/symbol.

than the target entropy $\mathbb{H}(P_{I,Q}) - 1$. On the other hand, the unconstraint bits that are mapped to the signs of the symbols induce no rate losses in the PAS scheme. Putting them together and assembling I, Q part, we know that the number of bits per symbol is $2(R_{dm} + 1)$, which normalizing with $2m$ leads to the shaping set rate $R_{ss} = \frac{R_{dm}+1}{m}$. Consequently, the total rate loss (defined as bits/symbol) of the QAM symbols induced by DM is:

$$\begin{aligned}\Delta_{PS} &= 2\mathbb{H}\left(P_I\right) - 2(R_{dm} + 1) \\ &= \mathbb{H}\left(X\right) - 2(R_{dm} + 1).\end{aligned} \tag{2}$$

When applying an FEC scheme, certain amount of overhead is induced. Let R_{fec} denote the FEC code rate. As shown in Fig. 1, for the real parts of symbol sequence of length n, the amplitudes are mapped to a sequence of $(m-1)n$ bits to apply the FEC scheme, and γn bits are used to encode actual data in the signs. A subsequence of length $(1-\gamma)n$ in the signs are used for FEC redundancy. By symmetry, the same relation holds for the imaginary part. Therefore, the FEC code rate satisfies the following identity:

$$R_{fec} = \frac{1}{1 + OH_{fec}} = \frac{m - 1 + \gamma}{m}. \tag{3}$$

Note that since the FEC code rate minimum threshold is $\frac{(m-1)}{m}$ [23], the FEC overhead is no more than $\frac{1}{m-1}$. (e.g., OH of PS 256QAM (m = 4) should be <33%). Therefore, for 64-QAM, we sweep the FEC OH from 5% to 30% considering practical implementation in the analysis in Section B and C.

The spectral efficiency of the entire system can then be written as:

$$SE = 2(R_{dm} + \gamma) = (R_{ss} - (1 - R_{fec})) \cdot \log_2 M. \tag{4}$$

The cost from the FEC component is subtracted from the PS rate in the final expression of spectral efficiency.

To compare with the target entropy, the spectral efficiency can be further written as:

$$SE = \mathbb{H}\left(X\right) - \Delta_{PS} - (1 - R_{fec})\log_2 M. \tag{5}$$

Fig. 2(a) shows the relationship between the term $(1 - R_{fec})\log_2 M$ (defined as the FEC redundancy in bits/symbol) and FEC OH for different QAM levels. Higher QAM order has a higher scaling factor of $(1 - R_{fec})$.

In above expression, two terms are separately subtracted from the entropy, standing for CCDM rate loss and FEC cost, respectively. With a sufficient symbol length, the rate loss from DM can be very small [22]. The second part is $(1 - R_{fec}) \cdot \log_2 M$, which corresponds to the redundancy from FEC encoding.

The information rate excluding the FEC redundancy of an upconverted PS-OFDM signal is defined as [24]:

$$R_{\text{info}} = \overline{SE} \cdot W \cdot \frac{1}{1 + r_{CP}}, \tag{6}$$

where $\overline{SE}$is the averaged SE of all the subcarriers, W is signal bandwidth and$r_{CP}$is the cyclic prefix (CP) ratio. On the other hand, the constellation entropy rate without considering the FEC OH is:

$$R_{raw} = \overline{\mathbb{H}\left(X\right)} \cdot W \cdot \frac{1}{1 + r_{CP}}, \tag{7}$$

where $\overline{\mathbb{H}(X)}$ is the averaged constellation entropy of all the subcarriers. When $r_{CP} = 0$, (6) and (7) correspond to the expression of a single carrier system: $R_{\text{info}} = SE \cdot W$ and $R_{raw} = \mathbb{H}(X) \cdot W$. Note that the CP ratio is usually fixed in experiment, thus we need to adjust target entropy $\mathbb{H}(X)$, FEC

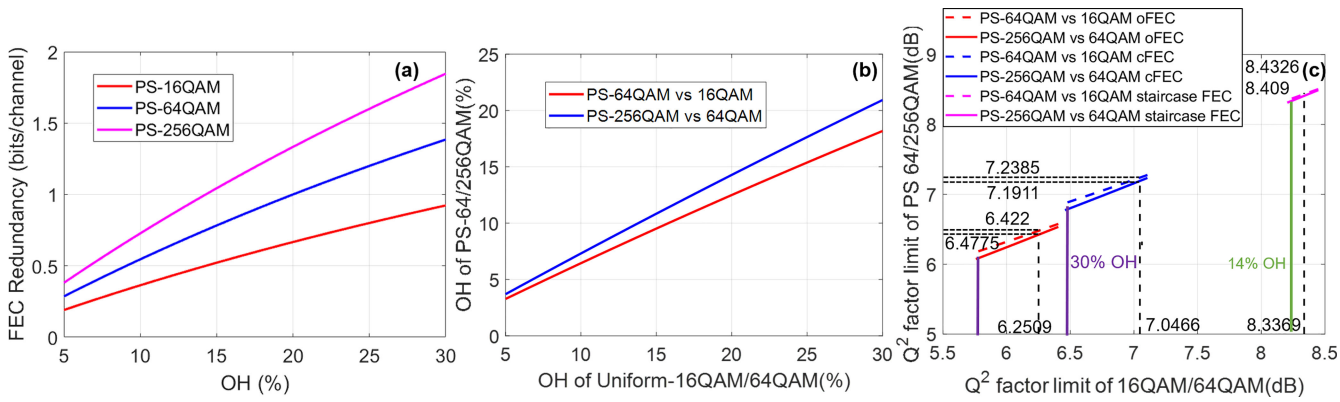

Fig. 2. (a) Fixed information rate case: FEC redundancy of information rate versus FEC OH of PS-16/64/256 QAM. Fixed information rate case: (b) Required FEC OH of PS-2^{2m+2} QAM versus 2^{2m} QAM OH to achieve the same SE (PS-64QAM 4.0 vs 16QAM and PS-256QAM 6.0 vs 64QAM). (c) Pre-FEC Q^2 factor limit of PS-2^{2m+2} QAM versus that of 2^{2m} QAM to achieve the same SE (PS-64QAM 4.0 vs 16QAM and PS-256QAM 6.0 vs 64QAM) with oFEC, cFEC, and staircase FEC respect to the post-FEC BER of 10^{-15}.

code rate R_{fec} and signal bandwidth W to ensure the same information rate for a fair comparison.

B. PS $2^{2m+2}QAM$ Versus Uniform $2^{2m}QAM$

Among recent experimental demonstrations, one of the common settings is to apply a PS signal with a higher-order QAM to compare with a uniform lower-order QAM. For instance, some research work compared the PS 64 QAM with Uniform 16 QAM or PS 256 QAM with Uniform 64 QAM [7].

To enable a PS signal with higher order modulation format, digital to analog converter (DAC) with higher resolution or effective number of bits (ENOB) is required. This method brings advantages regarding systems with limited analog bandwidth. In our comparison, we fix the bandwidth as W and spectral efficiency as SE for both PS and uniform signals.

To ensure the same SE in a bandwidth-limited system, with negligible rate loss from DM, we can apply Eq. (5) to both PS signal and Uniform signal, and make them equivalent:

$$\begin{aligned}&\mathbb{H}\left(X_{ps}\right)-\left(1-R_{fec}^{PS}\right)\log_2 2^{2m+2}\\&=\mathbb{H}\left(X_U\right)-\left(1-R_{fec}^{U}\right)\log_2 2^{2m}.\end{aligned}\tag{8}$$

The left side is the information rate of PS $2^{2m+2}QAM$ and right side is information rate of Uniform $2^{2m}QAM$. By fine tuning $\mathbb{H}(X)$ and R_{fec} on both sides of the equation, it can achieve the same SE as the Uniform signal. The choice of the target the entropy and FEC code rate depends on the experimental condition.

However, changing the FEC code rate for PS signal according to Eq. (8) also changes the pre-FEC BER thresholds. With the same post-FEC requirement, the Pre-FEC Q^2 factor or BER changes with the FEC code rate as shown below [25]:

$$\begin{aligned}Q^2(\mathrm{dB})=&\,20\log_{10}\left[erfc^{-1}\left(2\cdot BER_{post-FEC}\right)\right]\\&+10\log_{10}\left(R_{fec}\right)-NCG(\mathrm{dB})+20\log_{10}\sqrt{2}.\end{aligned}\tag{9}$$

Where $Q^2(\mathrm{dB})=20\log_{10}(\sqrt{2}\mathrm{erfc}^{-1}(2\mathrm{BER}_{\mathrm{pre-FEC}}))$. NCG corresponds to the net coding gain of the specific FEC with a certain OH implemented in experiment.

1) Modified Target Entropy Without Bandwidth Scaling: Many existing standards define the FEC scheme and pre-FEC thresholds as fixed. (e.g., in 400ZR OIF standard, 1.25e-2 pre-FEC BER with cFEC is utilized.). It is also very costly to perform rate adaptation in hardware implementation. Thus, here we discussed the case considering a fixed pre-FEC threshold for both PS signal and Uniform signal. In this case, according to Eq. (9), the FEC code rate of PS signal is the same as that of the Uniform signal, $R_{fec}^{PS}=R_{fec}^{U}=R_{fec}$ or $OH_{PS}=OH_U$. Then we can simplify Eq. (9) as:

$$\begin{aligned}\mathbb{H}\left(X_{PS}\right)-\mathbb{H}\left(X_U\right)&=\mathbb{H}\left(X_{PS}\right)-2m\\&=2(1-R_{fec}).\end{aligned}\tag{10}$$

The equation indicates that the target the entropy of the PS $2^{2m+2}QAM$ signal should be slightly higher than the entropy of Uniform$2^{2m}QAM$ signal. For instance, if the Uniform 16-QAM signal utilizes a FEC OH (OH_U) of 25%, the target entropy of the PS-64QAM signal should be increased to $4+(1-\frac{1}{1+OH_U})\cdot 2=4.4$ bits/symbol rather than 4 bits/symbol. PS 2^{2m+4} QAM vs 2^{2m} QAM would induce more significant OH penalty, but this case is not widely utilized in experimental demonstration.

2) Modified Pre-FEC Metrics Without Bandwidth Scaling: Note that some prior research works ignore the FEC redundancy discrepancy and perform the comparison with the same constellation entropy rate as shown in Eq. (7) [7]. For instance, to compare with a Uniform 16QAM signal, the PS 64QAM signal utilizes the entropy equals to 4 bits/symbol with the same bandwidth $(\mathbb{H}(X_{PS})=\mathbb{H}(X_U))$. The performance gain using this method might be over-estimated with the same Pre-FEC threshold. As discussed in part B.1, the assigned entropy for PS signal is smaller than that in Eq. (10) and reduces the information rate.

In the following analysis, we investigate the modified Pre-FEC Q^2 threshold for a fair comparison in case of $\mathbb{H}(X_{PS})=$

$\mathbb{H}(X_U) = 2m$. We can re-write Eq. (8) as:

$$(1 - R_{fec}^{PS})\log_2 2^{2m+2} = (1 - R_{fec}^{U})\log_2 2^{2m}. \quad (11)$$

To make the SE equal based on Eq. (11), the FEC overhead in uniform and PS signals satisfy the following equation (as shown in Fig. 2(b)):

$$OH_{PS} = \frac{1}{1 - (OH_U/(OH_U + 1)) \cdot m/(m+1)} - 1. \quad (12)$$

In Fig. 2(b), larger *m* exhibits a sharper slope of curves (closer to the slope of a 45° curve), which indicates that larger *m* will induce less OH difference in this application case. To further characterize the system performance in the experimental investigation, we plot the Pre-FEC Q^2 factor of PS 2^{2m+2} QAM versus Uniform 2^{2m} QAM in different cases of the three above-mentioned standard FECs. The above-mentioned Pre-FEC Q^2 factor thresholds (Q_1, Q_2, Q_3, Q_4) are utilized for the Uniform 2^{2m} QAM signal. Using Eq. (9), Eq. (11) and the post-FEC BER of 10^{-15}, we obtain the modified Pre-FEC Q^2 factor for PS 2^{2m+2} QAM.

In Fig. 2(c), three dashed lines correspond to the relationship between Q^2 factor limit of PS-64QAM and that of the 16QAM in case of $\mathbb{H}(X_{ps}) = \mathbb{H}(X_U)$ and the same SE. The three solid curves represent the case of PS-256 QAM versus 64QAM. The curves of the same color represent the same FEC scheme with OH varying around the OH defined by standards. For Staircase FEC, we sweep the OH value of Uniform signal from 4% to 10%. For oFEC and cFEc, we sweep the OH of Uniform signal from 12% to 30%. Three pre-FEC Q^2 limits are presented. We also plot the purple/green lines of the pre-FEC Q^2 limit at 30%/14% OH as the trend reference line. First, we can observe that for the same FEC scheme, more FEC OH would induce more Q^2 difference. The case of PS-64 QAM versus 16QAM exhibits more Q^2 factor difference compared with PS-256QAM vs 64QAM, given the same increase in FEC OH. However, with the decrease in the FEC OH (the increase in the requirement for pre-FEC Q^2 factor), the dashed curve and solid curve converges, which indicates that the Q^2 factor difference from different values of *m* will be closer to less FEC OH. We can also observe that, the modified pre-FEC Q^2 factor limit for PS-64QAM or 256QAM is more stringent than the Uniform signals For instance, $Q_3 = 8.3369$ dB for Uniform 16-QAM, while the Q^2 limit in Staircase FEC for PS-64QAM and PS-256QAM are 8.4326 dB and 8.409 dB, respectively. Moreover, a more powerful FEC with higher NCG shows less pre-FEC Q^2 factor thresholds difference due to less required OH.

As for NGMI, it has been proven to be a better predictor of soft-decision (SD) FEC [26]. It is obtained through generalized mutual information (GMI) by [16], [26]:

$$NGMI = 1 - (\mathbb{H}(X) - GMI)/\log_2 M, \quad (13)$$

where GMI is calculated from the log-likelihood ratios, $\mathbb{H}(X)$ is the constellation entropy and *M* is the QAM order of PS signal. GMI is an achievable information rate in the a bit-interleaved coded modulation (BICM)-AWGN auxiliary channel with deal binary FEC and bitwise decoding [26].

Considering the practical FEC always have gaps to the theoretical bound, we can re-write NGMI as $NGMI = R_{fec} + \Delta(\Delta \geq 0)$ [26]. According to Eq. (11), to ensure the same SE or GMI, with ideal FEC assumption ($\Delta = 0$) [26], we can modify the NGMI threshold for PS signal as:

$$NGMI_{PS} = \frac{NGMI_U \cdot 2m + 2}{2m + 2} = \frac{NGMI_U \cdot \mathbb{H}(X_U) + 2}{\mathbb{H}(X_U) + 2}. \quad (14)$$

Note that in practice, for a specific FEC, the discrepancy between NGMI and FEC code rate Δ should be considered when making the information rate equal.

3) Additional Required Bandwidth: On the other hand, we can also allocate more bandwidth without modifying the pre-FEC metrics in the case where $\mathbb{H}(X_{PS}) = \mathbb{H}(X_U)$and PS signal is used with higher order QAM. The bandwidth of PS signal can be obtained from Eq. (6) and (11) with $R_{fec}^{PS} = R_{fec}^{U} = R_{fec}$:

$$\mathrm{W}_{PS} = \frac{\mathbb{H}(X_U) - (1 - R_{fec}) \cdot 2m}{\mathbb{H}(X_{PS}) - (1 - R_{fec}) \cdot (2m + 2)} \cdot W_U = \frac{R_{fec} \cdot \mathbb{H}(X_U)}{R_{fec} \cdot (\mathbb{H}(X_U) + 2) - 2} \cdot W_U, \quad (15)$$

where W_{PS} is the bandwidth of PS signal and $\mathrm{W}_U$is the bandwidth of Uniform signal. The required additional bandwidth of PS signal is $(\mathrm{W}_{PS} - \mathrm{W}_U)/\mathrm{W}_U$, as shown in the blue and orange curves of Fig. 3(a).

C. PS $2^{2m}QAM$ Versus Uniform $2^{2m}QAM$

Different from the method in Section B, another setting that is commonly utilized in experiment is to apply a PS signal with the same order QAM with slightly lower entropy. For instance, some research work compared the PS 64 QAM with Uniform 64 QAM using slightly lower target entropy equals to 5.75 bits/symbol or 5.5 bits/symbol [1]–[4], [11], [27]. The target entropy of the PS $2^{2m}QAM$ cannot exceed that of Uniform $2^{2m}QAM$:$\mathbb{H}(X_{PS}) \leq \mathbb{H}(X_U) = 2m$. Therefore, this method puts less requirement on DAC resolution while needs FEC code rate adjustment or more analog bandwidth to ensure the same information rate.

1) Additional Required Bandwidth: Some research works also scale the PS signal bandwidth with a constant factor equals to $\frac{\mathbb{H}(X_U)}{\mathbb{H}(X_{PS})}$, which yields the same constellation entropy rate. However, to ensure the same information rate, the bandwidth scaling factor is not a constant value anymore. With $R_{fec}^{PS} = R_{fec}^{U} = R_{fec}$ and by making the information rates of PS signal and Uniform signal equal, we obtain:

$$\mathrm{W}_{PS} = \frac{\mathbb{H}(X_U) - (1 - R_{fec}) \cdot 2m}{\mathbb{H}(X_{PS}) - (1 - R_{fec}) \cdot 2m} \cdot W_U = \frac{\mathbb{H}(X_U) - (1 - R_{fec}) \cdot \mathbb{H}(X_U)}{\mathbb{H}(X_{PS}) - (1 - R_{fec}) \cdot \mathbb{H}(X_U)} \cdot W_U. \quad (16)$$

Fig. 3(a) and (b) present the modified required additional bandwidth in Eq. (11) of PS 64-QAM vs Uniform 64-QAM and PS-256QAM vs Uniform 256QAM, respectively. Dashed lines

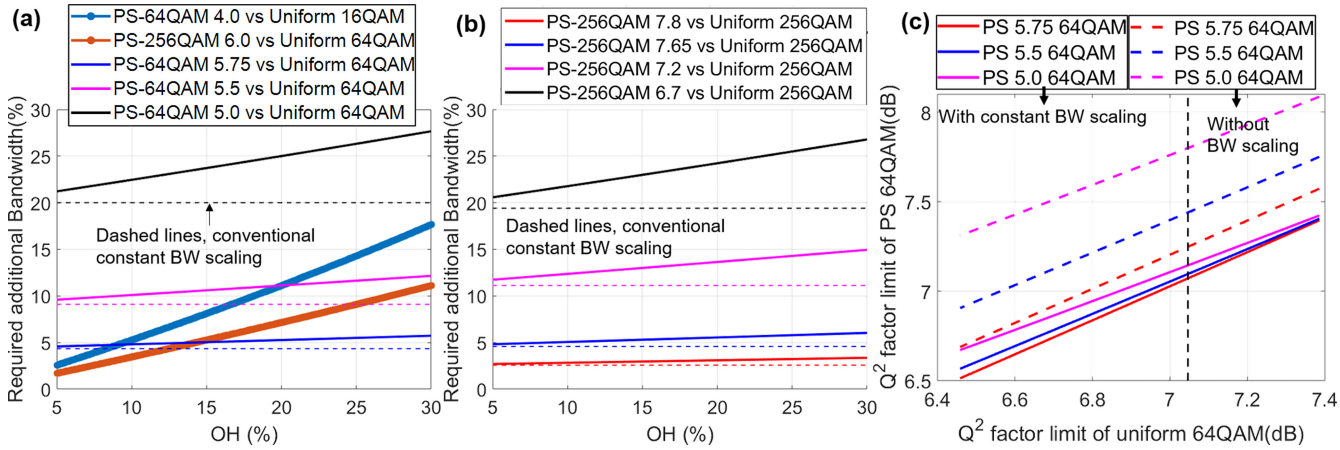

Fig. 3. With same information rate, required additional bandwidth versus FEC OH in case of (a) PS 64QAM 4.0 vs Uniform 16QAM, PS 256QAM 6.0 vs Uniform 64QAM, and PS 64QAM vs Uniform 64QAM;(b) PS 256QAM vs Uniform 256QAM. Reference (dashed) lines with conventional constant bandwidth scaling factor $\mathbb{H}(X_U)/\mathbb{H}(X_{PS})$ are provided. (c) Pre-FEC Q^2 factor limit of PS-64 QAM versus that of Uniform 64QAM to achieve the same information rate. Dashed lines are the PS signals without bandwidth scaling and solid lines are the PS signals with $\mathbb{H}(X_U)/\mathbb{H}(X_{PS})$ constant bandwidth scaling.

correspond to conventional way of scaling the bandwidth of PS signal using $\frac{\mathbb{H}(X_U)}{\mathbb{H}(X_{PS})}$ scaling factor, which is independent of the FEC OH. The smaller $\mathbb{H}(X_{PS})$ is, the more additional bandwidth is required and the system performance is more sensitive to the OH variation. With larger $\mathbb{H}(X_{PS})$ (e.g., 7.8 and 7.65), the required bandwidth remains nearly flat and is very close to the conventional required bandwidth (dashed lines) as shown in Fig. 3(b). Therefore, the conventional way of bandwidth scaling ignores the additional penalty to the effective data rate, especially for smaller $\mathbb{H}(X_{PS})$. However, in practice, dynamic bandwidth adjustment would be very costly since the allocated bandwidth is dependent on the FEC OH utilized in the system.

2) Modified Pre-FEC Metrics Without Bandwidth Scaling: In this part, we still consider ensuring the same SE without requiring more analog bandwidth. Making the SE of both sides equal, we obtain:

$$\begin{aligned}&\mathbb{H}\left(X_{PS}\right)-(1-R_{fec}^{PS})\log_2 2^{2m}\\&\quad=\mathbb{H}\left(X_U\right)-(1-R_{fec}^{U})\log_2 2^{2m}\\&\mathbb{H}\left(X_{PS}\right)-(1-R_{fec}^{PS})\mathbb{H}\left(X_U\right)\\&\quad=\mathbb{H}\left(X_U\right)-(1-R_{fec}^{U})\mathbb{H}\left(X_U\right)\\&R_{fec}^{ps}=1+R_{fec}^{U}-\frac{\mathbb{H}\left(X_{PS}\right)}{\mathbb{H}\left(X_U\right)}.\end{aligned}\tag{17}$$

Given R_{fec}^{ps} obtained from Eq. (13), we can obtain the modified Pre-FEC Q^2 factor based on Eq. (9).

According to Eq. (13), where $M=2^{2m}$ for both PS signal and Uniform signal, with ideal FEC and the same GMI or SE, we can obtain the modified NGMI threshold as:

$$NGMI_{PS}=1+NGMI_U-\frac{\mathbb{H}\left(X_{PS}\right)}{\mathbb{H}\left(X_U\right)}.\tag{18}$$

3) Modified Pre-FEC Metrics With Constant Bandwidth Scaling: On the other hand, to avoid dynamic bandwidth scaling, we can also keep the constant bandwidth scaling ($\mathrm{W}_{PS}=\frac{\mathbb{H}(X_U)}{\mathbb{H}(X_{PS})}\cdot W_U$, same constellation entropy rate) and modify the Pre-FEC metrics. Given $\mathrm{W}_{PS}==\frac{\mathbb{H}(X_U)}{\mathbb{H}(X_{PS})}\cdot W_U$ and the FEC code rate of Uniform signal R_{fec}^{U}, making the information rate of PS signal and Uniform signal equal, we obtain:

$$R_{fec}^{PS}=\frac{\left(R_{fec}^{U}-1\right)\cdot\mathbb{H}\left(X_{PS}\right)}{\mathbb{H}\left(X_U\right)}+1.\tag{19}$$

Combining Eq. (20) and Eq. (9) together, we can get the modified pre-FEC Q^2 factor with respect to a Q^2 factor limit of uniform signal and an NCG.

Fig. 3(c) presents an example of the modified pre-FEC Q^2 factor obtained from Eq. (20) and Eq. (18) with different $\mathbb{H}(X_{PS})$: Pre-FEC Q^2 threshold of PS 64QAM as a function of Pre-FEC Q^2 threshold of Uniform 64QAM. The figure is based on the cFEC (7.0466 dB Q^2 limit with OH 14.8% is also presented in a dashed black line). Without bandwidth scaling, PS 64QAM signals with 5 bits/symbol and 5.5 bits/symbol yields a significantly higher Q^2 factor. Even if using 5.75 bits/symbol, the modified Q^2 factor is still higher than that of the PS 256 QAM with 6 bits/symbol as shown in Fig. 2(c). However, with the constant bandwidth scaling with a factor of $\frac{\mathbb{H}(X_U)}{\mathbb{H}(X_{PS})}$, the difference between the modified Q^2 factor of PS signal and Q^2 factor of Uniform signal becomes smaller. In the experimental demonstration, we will utilize a constant bandwidth scaling, which yields a fixed constellation entropy rate and a smaller Q^2 limit change.

Similarly, with ideal FEC assumptions, the modified NGMI threshold can be obtained from Eq. (20):

$$NGMI_{PS}=\frac{(NGMI_U-1)\cdot\mathbb{H}\left(X_{PS}\right)}{\mathbb{H}\left(X_U\right)}+1.\tag{20}$$

In conclusion, with the same constellation entropy rate, fixed NGMI threshold and Pre-FEC Q^2 factor are not sufficient to evaluate the system performance directly to ensure a fair comparison. Therefore, in the experimental analysis, we need to ensure the

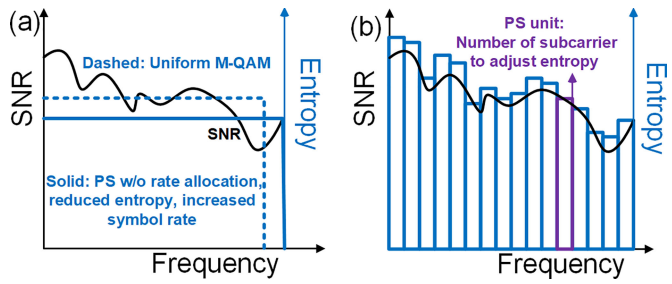

Fig. 4. Principle of operation of PS unit based entropy optimization. (Black curve: SNR over frequency; Blue curve: Entropy over frequency) (a) Uniform M-QAM or PS with no entropy allocation. (b) Entropy allocation based on PS unit.

same information by adjusting FEC code rate, signal bandwidth, or Pre-FEC thresholds as indicated before.

D. PS Unit Size Based Entropy Allocation

In this paper, as shown in Fig. 4, we proposed an entropy allocation scheme in an OFDM based MMW system by grouping several subcarriers into one PS unit. Assuming the OFDM signal has *T* subcarriers in total, we first divide the data into *S* PS units with *T/S* subcarriers per unit. We then optimize the entropy per PS unit based on the pre-estimated SNR:

$$\mathbb{H}\left(X_j\right) = \frac{S \cdot \overline{\mathbb{H}\left(X\right)} \cdot \log_2\left(1 + SNR_j\right)}{\sum_{i=1}^{S} \log_2\left(1 + SNR_i\right)}, \tag{21}$$

where $j \in [1, S]$, $\mathbb{H}(X_j)$and SNR_j denote allocated entropy (bits/symbol) and pre-measured SNR of the j_{th} PS unit, respectively. $\overline{\mathbb{H}(X)}$is the target averaged entropy of the whole data frame. Eq. (19) ensures that the averaged entropy of the PS units corresponds to $\overline{\mathbb{H}(X)}$. Note that all the entropies are tuned by changing the λin the MB distribution.

In practice, the ratio *T/S* can be adjusted while balancing the enhanced power margins, symbol latency and processing complexity. A smaller *T/S* ratio usually yields a better performance since it is better at tracking SNR fluctuations. However, smaller *T/S* ratio might induce extra latency and complexity due to the fact that we should have sufficient symbol length to ensure low rate loss [22]. On the other hand, the measured SNR might fluctuate with small symbol number, which makes the entropy allocation not precise. A larger *T/S* ratio would average part of the SNR fluctuations but relax the symbol number per subcarrier and has less estimated SNR fluctuations. Generally, a larger *T/S* ratio is more suitable for a smoother channel frequency response. The proposed entropy allocation scheme reduces the data frame latency (less symbols per subcarrier can be utilized) with tunable margins. Moreover, since this scheme utilizes the same QAM order in each PS unit with a fixed average entropy, it doesn't require dynamic FEC code rate adjusting or bandwidth adjusting considering different QAM orders utilized in different PS units (this will change the $\log_2 M$term in Eq. (20), giving different weights for FEC code rate), which reduces the processing complexity. Since each PS unit has even frequency spacing, assuming negligible rate loss from DM, the overall spectral efficiency can be expressed as:

$$\begin{aligned}\overline{SE} &= \frac{1}{S}\sum_{i=1}^{S}\left(\mathbb{H}\left(X_i\right) - \left(1 - R_{fec}\right)\cdot \log_2 M\right) \\ &= \overline{\mathbb{H}\left(X\right)} - \left(1 - R_{fec}\right)\cdot \log_2 M. \end{aligned} \tag{22}$$

Eq. (20) indicates that this scheme utilizes fixed FEC code rate for every PS unit over frequency and ensures the same SE as single carrier systems. Therefore, we can evaluate the performance of the whole data frame using the above mentioned modified thresholds based on target $\overline{\mathbb{H}(X)}$and QAM order M.

III. Experimental Setup

Fig. 5 illustrates the experimental setup of the hybrid fiber/MMW radio fronthaul link. In the transmitter DSP, a pseudorandom binary sequence (PRBS) is splitted into two sequences (Sequence 1 and Sequence 2) and loaded to the PAS scheme. Sequence 1 is divided into several PS units with different allocated entropies and then loaded to CCDM block to generate the desired probability of symbol amplitude, which obeys the MB distribution. Sequence 2 and LDPC parity matrix is utilized to form the signs of the PAM symbols. In the experiment, LDPC code rate is fixed as 0.8 and the binary labeling follows Gray encoding. After QAM assembling, we perform OFDM modulation and up-conversion. To investigate different cases, we adjust the signal SNR at the transmitter by changing the bandwidth based on different QAM orders. For the Uniform 16QAM symbols, there are 400 subcarriers and each of them has 54 symbols with a subcarrier spacing of 1.5MHz. The Uniform 64QAM signal consists of 400 subcarriers and 400 × 54 symbols with a subcarrier spacing of 500 kHz while Uniform 256QAM signal consists of 200 subcarriers and 200 × 81 symbols with a subcarrier spacing of 500 kHz. The IFFT size is 2048 and 5.12% CP is inserted. The signal generated through a digital to analog converter (DAC) and amplified via a power amplifier (PA). Then it is loaded to a 10-GHz bandwidth Mach-Zehnder Modulator (MZM). After an optical boost of Erbium-doped fiber amplifier (EDFA), the signal is delivered through a 15-km single mode fiber (SMF) and then detected by a photodetector (PD). After O/E conversion, an electrical mixer directly upconverts the electrical signal to a 60-GHz MMW signal using an RF source and a frequency quadrupler. After 1.5-m MMW transmission, an envelope detector (ED) is utilized to detect and down-convert the signal. In the receiver DSP, after down-conversion, synchronization and equalization, and OFDM demodulation, we perform inverse QAM assembling, which separates the sign and magnitude of the I and Q signal. Then we perform FEC decoding and calculate the post-FEC performance. Note that NGMI, GMI, pre-FEC BER or Q^2 is calculated after inverse binary labeling and inverse β operation and before LDPC decoding.

VI. Experimental Results and Analysis

In the following experimental analysis, we will compare the three schemes in the MMW based fronthaul system: (i) uniform QAM distribution, (ii) PS QAM with the same entropy per

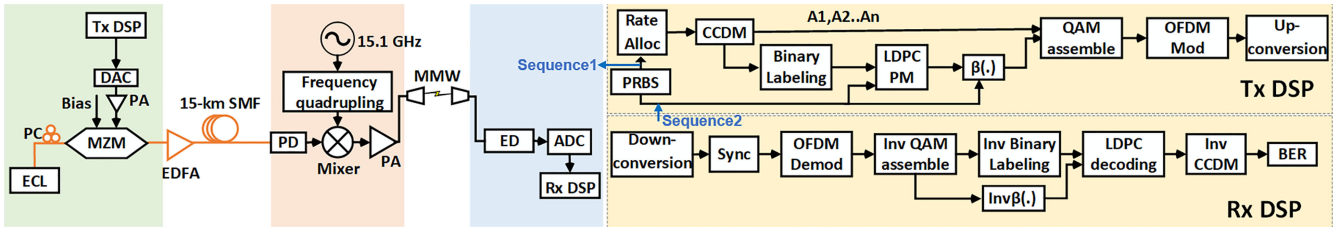

Fig. 5. Experimental setup and DSP diagram. DAC: digital to analog converter, ECL: external cavity laser, PA: power amplifier, EDFA: Erbium-doped fiber amplifier, PD: photodetector, ED: envelope detector, ADC: analog to digital converter, LDPC PM: LDPC parity matrix, sync: synchronization.

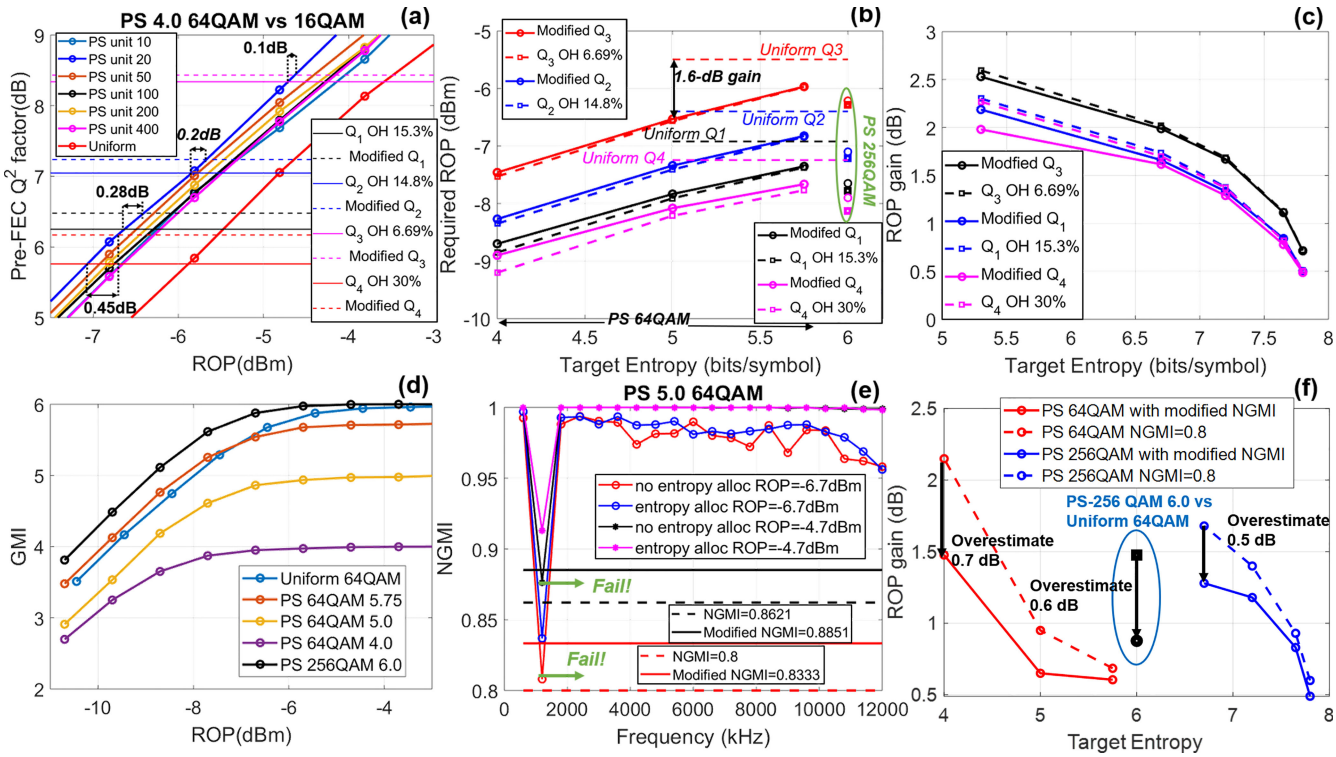

Fig. 6. Experimental results. (a) Pre-FEC Q^2 versus ROP using different sizes of PS units. (b) PS 64/256QAM versus Uniform 64 QAM. Required ROP versus target entropy with respect to standard Q thresholds and modified thresholds). (c) PS 256QAM versus Uniform 256QAM. ROP gain versus target entropy with standard Q (solid lines) and modified Q (dashed lines). (d) PS 64/256QAM versus Uniform 64QAM. GMI versus ROP. (e) NGMI of the PS 5.0 64QAM versus signal frequency with and without adaptive entropy allocation. (f) Red curves and blue square point: PS 64QAM/256QAM vs Uniform 64QAM. Blue curves: PS 256QAM vs Uniform 256QAM. ROP gain versus target entropy with respect to NGMI thresholds.

subcarrier in Fig. 4(a), and (iii) PS QAM aided adaptive rate allocation per PS unit as shown in Fig. 4(b). The analysis without consideration about FEC induced redundancy will be compared to investigate the over-estimation in the radio over fiber system.

A. *Pre-FEC Thresholds for Performance Evaluation*

In the experiment, we investigate the system performance with a fixed constellation entropy rate (scaling factor of bandwidth $=\frac{\mathbb{H}(X_U)}{\mathbb{H}(X_{PS})}$ as presented in Section B.2 and C.3). The above-mentioned modified Q^2 limit is utilized to investigate the system performance gain and over-estimation. The Pre-FEC thresholds of the PS signals are modified with respect to these thresholds of uniform signals (summarized in Table I) according to Eq. (9) in Section B, Eq. (11), (14) in Section B.2, and Eq. (19), (20) in Section C.3.

B. *Optimization of Subcarrier Number Per PS Unit*

Fig. 6(a) illustrates the sensitivity performance of PS unit size sweeping using PS-4.0 64QAM versus Uniform 16QAM with the same occupied bandwidth. The solid horizontal lines are reference lines of Q_1, Q_2, Q_3, Q_4 for Uniform 16QAM signal while the dashed horizontal lines are the modified reference lines of Q_1, Q_2, Q_3, Q_4 according to Eq. (17).

In the figure, we can observe the over-estimation of the gain using the gap between the two received optical power (ROP) values with respect to the threshold of the uniform the signal and

the modified version of the PS signal. With respect to Q_3 with a smaller FEC OH, the over-estimation is only 0.1 dB. However, the over-estimation changes to 0.2 dB, 0.28 dB and 0.45 dB regrading Q_2, Q_1, and Q_4 respectively. The trend indicates that utilizing higher FEC OH would induce a more significant over-estimation if not considering net bit rate.

Since the channel fluctuation is relatively flat, 400, 200, and 100 subcarriers per PS unit show similar performance. Note that 400 subcarriers per PS unit corresponds to a PS signal without entropy allocation. With fine adjustment, the sensitivity gain significantly increases with the decreasing of the number of subcarriers per PS unit block. However, when the PS unit size is reduced to 10 subcarriers, the performance becomes worse. This is due to insufficient symbol length per PS unit for an accurate distribution matching and SNR estimation fluctuations. The issue can be solved by increasing the symbols per subcarrier to ensure sufficient symbol length for PS but it will increase the frame duration and latency. In our system, the PS unit size 20 achieves up to 1.1-dB ROP gain. In the following experiment, we utilize the optimized PS unit size of 20 for PS signal. It is worth noting that, different systems have different channel responses and symbol durations, so the optimal PS unit size might change with the optimization procedure.

C. Sensitivity Performance

Fig. 6(b) and (c) present the performance gain with respect to different pre-FEC Q^2 thresholds. The gap between the same colored dashed curve and solid curve indicates the over-estimated gain. Fig. 6(b) illustrates the case of PS 64QAM and PS 6.0 256QAM versus Uniform 64QAM and Fig. 6(c) demonstrates the case of PS 256QAM versus Uniform 256QAM. The dashed lines correspond to the over-estimated case while the solid curves present the required ROP with modified Q. Same color stands for the same pre-FEC Q factor of uniform signal. In Fig. 6(b), four horizontal reference lines are also presented, which present the required ROP of Uniform 64QAM signal with respect to the four Q factor limits. The square points in Fig. 6(b) are the over-estimated cases of PS 6.0 256QAM while circle points with the same color are the case using modified pre-FEC limits. Similar to the trend observed in Fig. 6(a), in both Fig. 6(b) and (c), the threshold with less FEC OH (e.g., Q_3) generates less gap between the same colored dashed curve and solid curve. Moreover, with the increasing of the target entropy, the gap between the dashed curve and solid curve becomes smaller. The largest over-estimation gap can be observed is 0.35 dB.

In Fig. 6(b), compared with Uniform 64QAM case regarding Q_3 threshold, with slightly increased bandwidth (4.34%), PS-5.75 shows up to 0.6-dB sensitivity gain. On the other hand, PS-4.0 shows highest ROP gain (up to 2 dB) but needs 20% extra bandwidth. Regarding the threshold with less FEC OH such as Q_3, the gain from PS 6.0 256QAM is smaller than PS 5.0 64QAM. However, with respect to higher FEC OH and smaller pre-FEC Q factor limit, the gain from PS 6.0 256QAM can achieve similar gain compared to PS 5.0 64QAM. Therefore, when the bandwidth is strictly limited, PS-6.0 256-QAM without any bandwidth scaling as indicated in the square box is the better choice. In Fig. 6(c), up to 2.5-dB sensitivity gain can be observed by using the proposed entropy allocation scheme. Compared with PS 64-QAM signal, higher order QAM can achieve more improvement by using PS. In practice, target entropy should be chosen considering performance gain, DAC resolution, and system extra bandwidth availability.

Fig. 6(d) depicts the GMI versus ROP in different cases. GMI is a good indicator of the achievable information rate (AIR) assuming ideal FEC. PS 6.0 256QAM achieves a better GMI compared with Uniform 64QAM. Moreover, we can observe that in the low ROP region (< -7.5 dBm), PS 5.75 64QAM has a better GMI in comparing to the uniform 64QAM thanks to the entropy allocation with increased Euclidean distance. When the ROP gets higher, uniform 64QAM outperforms PS 5.75 64QAM since the PS 5.75 64QAM tends to saturate when the received SNR is high.

Fig. 6(e) compares the performance of PS-5.0 64-QAM signal without entropy adaptation and with entropy adaptation, respectively. The results of ROP equals to -6.7 dBm and -4.7 dBm are presented. By allocating entropy dynamically using a PS unit of 20 subcarriers per block, the NGMI over frequency domain is optimized and becomes smoother. Moreover, the NGMI thresholds (dashed horizontal lines) of 0.8 and 0.8621, the corresponding modified NGMI for PS 5.0 64QAM signal (solid horizontal lines) from Eq. (18) are presented. If using the over-estimated NGMI thresholds as shown in dashed lines, both schemes with ROP of -6.7 dBm can pass 0.8 threshold while both with ROP of -4.7 dBm can pass 0.8621. However, by using the modified NGMI threshold, the scheme without entropy allocation will fail in some PS units (fall below the solid reference lines). Fig. 6(f) presents the ROP gain with NGMI threshold of the uniform signals equals to 0.8. The dashed curves present the over-estimated case while the solid curves correspond to the gain regarding modified NGMI. The two red curves are the case of PS 64QAM versus Uniform 64QAM while the two blue curves represent PS 256QAM vs Uniform 256QAM. The black square point and circle point indicate the ROP gain of PS 6.0 256QAM versus Uniform 64QAM without modified NGMI threshold and with NGMI threshold, respectively. PS 6.0 256QAM shows a 0.6-dB over-estimation of ROP gain, which makes it even worse than PS 5.0 64QAM. Up to 0.7-dB over-estimation and 0.5-dB over-estimation can be observed for PS 64QAM vs Uniform 64QAM and PS 256QAM vs Uniform 256QAM, respectively. In case of PS 5.75 64QAM and PS 7.8 256QAM, the overestimation is not significant.

V. Conclusion

In the experimental evaluation, it is important to make the information rate the same between different PS schemes for the purpose of a fair comparison. In this paper, different operation scenarios are discussed, including adjusting target entropy of PS signal, scaling the signal bandwidth, and changing the FEC code rate for preserving the same information rate. To investigate existing experimental scenarios with fixed contestellation entropy rate, the modified Pre-FEC thresholds of PS signal (obtained from the modification of FEC code rate) is proposed to provide an accurate performance evaluation.

We also propose an entropy allocation scheme, evaluate its sensitivity performance, and investigate over-estimation issue in a millimeter-wave based radio access system. The scheme relaxes the data frame length requirement by grouping several subcarriers into one PS unit for entropy tuning and reduce the processing complexity by avoiding dynamic FEC code rate adjusting. In our setup, optimal PS unit size with 20 subcarriers selected from PS unit size ranging from 10 subcarriers to 400 subcarriers can achieve a balance between the required sensitivity and processing complexity. PS unit size and target entropy tunability achieve a flexible adjustment of power margins gain from 0.6 dB to 2.5 dB with consideration for available analog bandwidth and DAC resolution. Moreover, we observe the over-estimation in the ROP gain up to 0.7 dB in the experiment. The mobile fronthaul system would result in a reduction of the information rate if one utilizes constellation entropy rate without considering FEC redundancy, especially in the case of lower target entropy and higher pre-FEC OH. Note that in practice, we can also modify the target entropy of PS signal (see, Section B.1) or scale the signal bandwidth (see, Section B.3 and C.1) without modifying the FEC code rate based on the hardware feasibility. With the proposed modified pre-FEC thresholds, our proposed entropy allocation shows significant performance gain while retaining the same information rate for analog RoF fronthaul.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for valuable input and suggestions that significantly improved this work.

REFERENCES

[1] L. Zhao, Y. Zhang, and W. Zhou, "Probabilistically shaped 64QAM OFDM signal transmission in a heterodyne coherent detection system," *Opt. Commun.*, vol. 434, pp. 175–179, 2019.

[2] B. Liu, X. Li, Y. Zhang, X. Xin, and J. Yu, "Probabilistic shaping for ROF system with heterodyne coherent detection," *Apl Photon.*, vol. 2, 2017, Art. no. 056104.

[3] H.-C. Chien, J. Yu, Y. Cai, B. Zhu, X. Xiao, Y. Xia, "Approaching Terabits per carrier metro-regional transmission using beyond-100GBd coherent optics with probabilistically shaped DP-64QAM modulation," *J. Lightw. Technol.*, vol. 37, pp. 1751–1755, Apr. 2019.

[4] A. Josten, B. Baeuerle, B. I. Bitachon, G. Stalder, D. Hillerkuss, and J. Leuthold, "400G probabilistic shaped PDM-64QAM synchronization in the frequency domain," *IEEE Photon. Technol. Lett.*, vol. 31, pp. 697–700, May 019.

[5] M. Kong *et al.*, "WDM transmission of 600G carriers over 5,600 km with probabilistically shaped 16QAM at 106 gbaud," in *Proc. Opt. Fiber Commun. Conf. Exhib.*, 2019, pp. 1–3.

[6] X. Hong, C. Fei, G. Zhang, and S. He, "Probabilistically shaped 256-QAM-OFDM transmission in underwater wireless optical communication system," in *Proc. 2019 Opt. Fiber Commun. Conf. Exhib.*, 2019, pp. 1–3.

[7] T. Sasai, A. Matsushita, M. Nakamura, S. Okamoto, F. Hamaoka, and Y. Kisaka, "Experimental analysis of laser phase noise tolerance of uniform 256QAM and probabilistically shaped 1024QAM," in *Proc. 2019 Opt. Fiber Commun. Conf. Exhib.*, 2019, pp. 1–3.

[8] K. Wu, J. Hea, Z. Zhou, and J. Shi, "Probabilistic amplitude shaping for a 64-QAM OFDM W-band RoF system," *IEEE Photon. Technol. Lett.*, vol. 31, no. 13, pp. 1076–1079, Jul. 2019.

[9] X. Li *et al.*, "1-Tb/s millimeter-wave signal wireless delivery at D-band," *J. Lightw. Technol.*, vol. 37, no. 1, pp. 196–204, Jan. 2019.

[10] M. Nakamura *et al.*, "104 Tbps/carrier probabilistically shaped PDM-64QAM WDM transmission over 240 km based on electrical spectrum synthesis," in *Proc. Opt. Fiber Commun. Conf.*, 2019, Paper M4I. 4.

[11] Y. Zhu *et al.*, "Spectrally-efficient single-carrier 400G transmission enabled by probabilistic shaping," in *Proc. Opt. Fiber Commun. Conf.*, 2017, Paper M3C. 1.

[12] T. Fehenberger, G. Böcherer, A. Alvarado, and N. Hanik, "LDPC coded modulation with probabilistic shaping for optical fiber systems," in *Proc. 2015 Opt. Fiber Commun. Conf. Exhib.*, 2015, pp. 1–3.

[13] Z. Cao, L. Zou, L. Chen, and J. Yu, "Impairment mitigation for a 60 GHz OFDM radio-over-fiber system through an adaptive modulation technique," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 3, no. 9, pp. 758–766, Sep. 2011.

[14] C. Xie *et al.*, "Achievable information rate enhancement of visible light communication using probabilistically shaped OFDM modulation," *Opt. Express*, vol. 26, pp. 367–375, 2018.

[15] R. Zhang *et al.*, "Joint optimization of processing complexity and rate allocation through entropy tunability for 64-/256-QAM based radio fronthauling with LDPC and PAS-OFDM," in *Proc. Opt. Fiber Commun. Conf.*, 2020, Paper M2F.2.

[16] J. Cho, "Balancing probabilistic shaping and forward error correction for optimal system performance," in *Proc. Opt. Fiber Commun. Conf.*, 2018, Paper M3C. 2.

[17] K. Wang, Y. Wei, M. Zhao, W. Zhou, and J. Yu, "140-Gb/s PS-256-QAM transmission in an OFDM system using kramers–kronig detection," *IEEE Photon. Technol. Lett.*, vol. 31, no. 17, pp. 1405–1408, Sep. 2019.

[18] J. Cho, X. Chen, S. Chandrasekhar, and P. Winzer, "On line rates, information rates, and spectral efficiencies in probabilistically shaped QAM systems," *Opt. Express*, vol. 26, pp. 9784–9791, 2018.

[19] G. Böcherer, P. Schulte, and F. Steiner, "Probabilistic shaping and forward error correction for fiber-optic communication systems," *J. Lightw. Technol.*, vol. 37, no. 2, pp. 230–244, Jan. 2019.

[20] F. Steiner, P. Schulte, and G. Bocherer, "Approaching waterfilling capacity of parallel channels by higher order modulation and probabilistic amplitude shaping," in *Proc. 52nd Annu. Conf. Inf. Sci. Syst*, 2018, pp. 1–6.

[21] Y. Cai *et al.*, "FPGA investigation on error-flare performance of a concatenated staircase and hamming FEC code for 400G inter-data center interconnect," *J. Lightw. Technol.*, vol. 37, pp. 188–195, Jan. 2019.

[22] P. Schulte and G. Böcherer, "Constant composition distribution matching," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 430–434, Jan. 2016.

[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] J. R. Barry, D. G. Messerschmitt, and E. A. Lee, *Digital Communication*. 3rd ed., Norwell, MA, USA: Kluwer Academic Publishers, 2003.

[25] T. Mizuochi, "Recent progress in forward error correction and its interplay with transmission impairments," *IEEE J. Sel. Topics Quantum Electron.*, vol. 12, no. 4, pp. 544–554, Jul./Aug. 2006.

[26] J. Cho, L. Schmalen, and P. J. Winzer, "Normalized generalized mutual information as a forward error correction threshold for probabilistically shaped QAM," in *Proc. 2017 Eur. Conf. Opt. Commun*, 2017, pp. 1–3.

[27] X. Li *et al.*, "132-Gb/s photonics-aided single-carrier wireless terahertz-wave signal transmission at 450GHz enabled by 64QAM modulation and probabilistic shaping," in *Proc. Opt. Fiber Commun. Conf*, 2019, Paper M4F.4.