Photonics-Assisted Millimeter-Wave Communication System Based on Low-Bit Gaussian Mixture Model Adaptive Vector Quantization

Yuancheng Cai [ID], *Member, IEEE*, Linghao Yue, Min Zhu [ID], *Member, IEEE*, Mingzheng Lei [ID], *Member, IEEE*, Jiao Zhang [ID], *Member, IEEE*, Bingchang Hua, Wei Luo, Yucong Zou, Liang Tian, Like Ma [ID], and Jianjun Yu [ID], *Fellow, IEEE*

Abstract—The quantization technique with a low-bit resolution can significantly reduce the cost and power consumption of analog-to-digital converter (ADC). It will play an important role in energy conservation and cost reduction for the incoming B5G millimeter-wave (MMW) communication systems. In this paper, we propose and demonstrated experimentally a low-bit Gaussian mixture model (GMM) based non-uniform adaptive vector quantization (AVQ) scheme for the low-cost intensity modulated envelope detection photonics-assisted 28 GHz MMW communication system for the first time. The principles of GMM-based one-dimensional adaptive scalar quantization (ASQ) and multi-dimensional AVQ are first introduced, and then are used to realize the low-bit non-uniform adaptive quantization for reducing the ADC bit resolution of MMW receiver. Furthermore, the performance of traditional uniform quantization, the present *K*-means and proposed GMM-based non-uniform ASQ/AVQ schemes are evaluated and compared in detail. Utilizing the proposed GMM-based AVQ scheme, the ADC quantization resolution in our MMW receiver can be reduced from 5 bits of the traditional uniform quantization to as low as 2 bits, without noticeable performance penalty. Moreover, as compared with the *K*-means-based quantization scheme, the MMW receiver enabled by GMM-based ASQ/AVQ scheme can save about half of the quantization time under similar performance. This is mainly because the clustering based on probability converges faster than the Euclidean distance, which significantly reduces the number of iterations required. Therefore, the GMM-based AVQ scheme is a promising solution to realize high performance ADCs with low-bit resolution for future MMW-enabled optical fiber wireless access networks.

Index Terms—Vector quantization, Gaussian mixture model, K-means clustering, low bit, MMW, optical fiber wireless access.

I. INTRODUCTION

WITH the explosive growth of a variety of emerging Internet services, such as virtual reality, digital twin, holographic projection and metaverse, the throughput and bandwidth of wireless communication networks increase dramatically [1]. The photonics-assisted millimeter-wave (MMW) communication system, which supports the large-capacity communication and wide-area coverage of MMW signal by means of radio-over-fiber architecture, has become the key technology in the upcoming beyond 5G and 6G era [2], [3]. However, the ever-increasing of throughput and bandwidth puts forward higher requirements for overall cost and power consumption of this MMW system, which mainly depend on the modulation and detection methods as well as signal quantization techniques, etc. In the photonics-assisted MMW communication system, simple intensity modulation with the high-order modulation format such as four-level pulse amplitude modulation (PAM4) and low-cost envelope detection based on the Schottky barrier diode are effective strategies to reduce its cost and power consumption [4]. Accordingly, the quantization technique has become one hotspot and the next breakthrough to realize energy conservation and cost reduction. In particular, under the scenario of massive machine type communication in future B5G MMW systems, the cost and power consumption of high-bit resolution analog-to-digital converters (ADCs) at the MMW receiver end for high-speed signal quantization is a vital challenge [5]. As is known, the cost and power consumption of ADC mainly depends on the quantization bit resolution when at the same sampling rate. Therefore, how to realize a low-bit quantization without inducing the obvious performance penalty is the key to solving the above problem.

Uniform quantization is the simplest quantization method, and it usually requires a high bit resolution to ensure adequate

Yuancheng Cai, Min Zhu, and Jiao Zhang are with National Mobile Communications Research Laboratory, Southeast University, Nanjing, Jiangsu 210096, China, and also with Purple Mountain Laboratories, Nanjing, Jiangsu 211111, China (e-mail: caiyuancheng@pmlabs.com.cn; minzhu@seu.edu.cn; jiaozhang@seu.edu.cn).

Linghao Yue and Wei Luo are with National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: yuelinghao@seu.edu.cn; 220211043@seu.edu.cn).

Mingzheng Lei, Bingchang Hua, Yucong Zou, and Liang Tian are with Purple Mountain Laboratories, Nanjing, Jiangsu 211111, China (e-mail: 2016010326@bupt.cn; huabingchang@pmlabs.com.cn; zouyucong@pmlabs.com.cn; tianliang@pmlabs.com.cn).

Like Ma is with Future Mobile Technology Lab, China Mobile Research Institute, Beijing 100053, China (e-mail: malike@chinamobile.com).

Jianjun Yu is with the School of Information Science and Technology, Fudan University, Shanghai 200433, China (e-mail: jianjun@fudan.edu.cn).

TABLE I
SUMMARY OF QUANTIZATION SCHEME IN FIBER OR WIRELESS COMMUNICATION SYSTEMS

Year & Reference	Object	Quantization scheme	QBs/Sa (lowest)	Methodology	Quantization level	Quantization scale	PDF estimation	Verification
2011 [6]	DAC & ADC	Uniform	3	Amplitude clipping	Fixed	Scalar	No	Simulation & Experiment
2019 [9]	DAC	Non-uniform	6	A-law & Lloyd	Fixed	Scalar	No	Experiment
2018 [10]	DAC	Non-uniform	3	A-law & Lloyd	Fixed	Scalar	No	Experiment
2018 [12]	DAC	Non-uniform	3	NPHE	Adaptive	Scalar	Yes	Simulation & Experiment
2017 [13]	DAC	Non-uniform	4	GLE	Adaptive	Scalar	Yes	Simulation
2017 [14]	DAC	Non-uniform	3	GLE	Adaptive	Scalar	Yes	Simulation & Experiment
2016 [15]	ADC	Non-uniform	2	IDM-MLSE[a]	Adaptive	Scalar	Yes	Simulation
2022 [16]	DAC	Non-uniform	3	GMM	Adaptive	Scalar	No	Simulation
2017 [17]	ADC	Non-uniform	5	Lloyd	Adaptive	Vector	No	Simulation
2019 [19]	ADC	Non-uniform	4	FSE[b]	Adaptive	Vector	No	Simulation
2018 [20]	ADC	Non-uniform	3	SOFM-NNC[c]	Adaptive	Vector	No	Experiment
2018 [21]	ADC	Non-uniform	3	K-means	Adaptive	Vector	No	Experiment
2021 [23]	ADC	Non-uniform	4	DCT[d] with K-means	Adaptive	Vector	No	Experiment
2019 [24]	ADC	Non-uniform	3	VLP[e] with K-means	Adaptive	Vector	No	Experiment
This work	ADC	Non-uniform	2	GMM	Adaptive	Vector	No	Experiment

[a]IDM-MLSE: Inclusive distortion model for maximum-likelihood sequence estimation; [b]FSE: Fast statistical estimation; [c]SOFM-NNC: Self-organizing feature map neural network clustering; [d]DCT: Discrete cosine transform; [e]VLP: Vector linear prediction.

quantization performance. Two approaches including amplitude clipping [6], [7] and peak-to-average power ratio (PAPR) mitigation [8] can be used to reduce the need for quantization bits. However, both of the two approaches have a limited reduction in quantization bit resolution due to the presence of extra noise induced by the clipping or PAPR reduction. On the other hand, non-uniform quantization is another effective method to reduce the required number of quantization bits. To date, some non-uniform quantization schemes with low-bit resolution have been presented, achieving the improved quantization performance to some extent [9], [10], [11], [12], [13], [14], [15]. To be specific, these non-uniform quantization schemes mainly include the A-law quantization [9], [10], Lloyd quantization [11], non-parametric histogram estimation (NPHE) [12], Gaussian-like estimation (GLE) [13], [14] and other novel quantization schemes [15]. Nevertheless, these above-mentioned quantization schemes either have the fixed quantization levels which cannot be adjusted adaptively according to the actual distribution of input signal, or need to extra estimation the probability density function (PDF) of the signal, which increases the complexity of the quantization operation. Fortunately, the novel low-bit adaptive quantization schemes without PDF estimation have also proposed and demonstrated [16], [17], [18], [19], [20], [21], [22], [23], [24]. Among them, the K-means-based non-uniform quantization is a commendable scheme and has obtained extensive attention in recent years. Although it has been proved the K-means quantization method can provide excellent quantization performance [21], the hard decision based on Euclidean distance clustering demands a large number of iterations which would bring high complexity [16].

In this paper, for the first time, we present and experimentally verify an adaptive vector quantization (AVQ) scheme based on the Gaussian mixture model (GMM) clustering algorithm for the photonics-assisted MMW communication system. The performance of different quantization schemes, including traditional uniform quantization, the present K-means and proposed GMM-based one-dimensional adaptive scalar quantization (ASQ) and two-dimensional AVQ, are evaluated in detail. The detailed comparison between this work and previous reports is summarized in Table I. We list the key advantages of the proposed photonics-assisted 28 GHz MMW system based on the GMM AVQ scheme in the following. It mainly includes three aspects: (1) supports vector quantization and the quantization levels can be adjusted adaptively according to the actual distribution of MMW signal without PDF estimation. (2) The quantization bit resolution as low as 2 bits can be realized in our proposed photonics-assisted MMW communication system enabled by the GMM-based AVQ scheme, which can significantly reduce the cost and power consumption of the ADCs in MMW receiver end, as compared with traditional uniform quantization scheme. (3) The proposed GMM-based ASQ and AVQ schemes can achieve similar performance as that of the K-means approach, whereas the running time can be reduced by about half under the same convergence threshold.

This paper is organized in four Sections. Section II introduces the principles of the K-mean and GMM-based AVQ scheme, respectively. Experimental verification and result discussions are shown in Section III. In the last Section, the main conclusions are given eventually.

II. PRINCIPLES OF K-MEANS AND GMM-BASED AVQ SCHEME

A. Principle of Vector Quantization

Suppose the target dataset contains a total of N_0 scalar samples to be quantified. Generally speaking, the scalar samples are quantified on the one-dimensional scale. However, in order to reduce the quantization distortion and improve the quantization performance, the vector quantization is proposed and performed on the multi-dimensional scale. At this case, the above N_0 scalar samples should be regrouped into N vector samples firstly. Here, we use d to present the dimension of vector quantization. For the ASQ scheme, the scalar dataset is original $\{x_1, x_2, \ldots, x_N\}$, where $N = N_0$. Instead, for the AVQ scheme with the same numbers of total samples, the vector dataset is given as $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\}$, where $N = \lceil N_0/d \rceil$ and the symbol $\lceil A \rceil$ stands for rounding up A. The constructed i-th vector sample can be expressed as

$$\boldsymbol{x}_i = \left(x_{(i-1)d+1}, x_{(i-1)d+2}, \ldots, x_{id}\right), \quad (1 \le i \le N) \quad (1)$$

Suppose the total number of quantization levels is M, the $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_M\}$ represents the codebook of vector quantization, and $\boldsymbol{c}_j$ is the j-th vector quantization level (i.e., codeword) in the codebook. Each vector sample $\boldsymbol{x}_i$ and codeword $\boldsymbol{c}_j$ are all composed of d-dimensional scalars. The process of vector quantization is to match each vector sample in the dataset $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\}$ to one of codewords in the codebook $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_M\}$ on the d-dimensional scale.

The vector quantization is applicable to both real-valued and complex-valued signals. In theory, the original scalar samples can be acquired just from the real/imaginary part of the signal, or from a joint combination of real and imaginary parts of the complex-valued signal. However, in practice, the former is the preferred solution. Because the scalars in each vector sample $\boldsymbol{x}_i$ have the correlation due to symbol up-sampling, while the latter case does not. As is known, the samples up-sampled from the same symbol are numerically close to each other. Leveraging on this, the obtained vector samples which reconstructed from the above-mentioned scalar samples are more centralized along positive proportion distribution (e.g., $y = x$ for two-dimensional vector quantization) [21]. Benefitting from this correlation, multi-dimensional vector quantization can further reduce the quantization distortion, as compared with traditional uniform quantization scheme.

When talking about the cost and power consumption of the quantizer, it is undoubtedly related to the parameter of quantization bits per sample (*QBs/Sa*). Notably, M quantization levels correspond to $\log_2(M)$ quantization bits for both ASQ and AVQ schemes. For the one-dimensional ASQ scheme, the *QBs/Sa* corresponds to $\log_2(M)$. Nevertheless, since each vector $\boldsymbol{x}_i$ contains d scalar samples for the d-dimensional AVQ scheme, thus the *QBs/Sa* is $\log_2(M)/d$. On the other hand, in order to achieve adaptive quantization, the quantization levels in the codebook should be adjusted adaptively according to the different distribution of input signal. In the following, the AVQ based on K-means and GMM clustering algorithms are presented to update the codebook adaptively.

B. Principle of K-means-Based AVQ Scheme

Based on the idea of hard decision, the K-means clustering algorithm adopts Euclidean distance to divide the dataset with N objects into M clusters. Firstly, the initial centroids $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_M\}$ of the M clusters can be initialized randomly. Secondly, according to the nearest neighbor principle, all samples of the input dataset $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\}$ are divided into corresponding centroid of the M clusters in turn. Thirdly, update the centroid for each cluster by taking the average operation according to

$$\boldsymbol{c}'_j = \frac{1}{N} \sum_{\boldsymbol{x}_i \in \boldsymbol{c}_j} \boldsymbol{x}_i, \quad (1 \le i \le N, 1 \le j \le M) \quad (2)$$

The quantization distortion is assessed through the sum of square differences between each sample and its quantized code-word, which can be expressed as

$$D_P = \sum_{j=1}^{M} \sum_{\boldsymbol{x}_i \in \boldsymbol{c}'_j} \left\| \boldsymbol{x}_i - \boldsymbol{c}'_j \right\|^2, \quad (1 \le i \le N, P \ge 1) \quad (3)$$

where P represents the number of iteration. A convergence threshold can be preset. The iteration operation is stopped until the ratio of the difference between the two adjacent quantization distortions, i.e., $(D_{P+1} - D_P)/D_P$, is less than the convergence threshold. To a certain extent, multiple iterations can avoid the problem of K-means clustering convergence to a local optimal solution.

Utilizing the above clustering algorithm, the codebook $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_M\}$ of K-means non-uniform quantization can be updated adaptively according to the distribution of input signal, without calculating its PDF. The concept diagrams of traditional uniform quantization, and K-means-based one-dimensional ($d = 1$) ASQ as well as two-dimensional ($d = 2$) AVQ schemes are shown in Fig. 1. Different from the traditional uniform quantization, the quantization levels of K-means-based ASQ are unevenly distributed on the one-dimensional scale, which implies more quantization levels can be allocated for the samples in dense area, contributing to a lower quantization distortion. For the K-means-based AVQ scheme, each two-dimensional vector $\boldsymbol{x}_i$ falls into the corresponding Voronoi cell formed by Voronoi boundaries, and will be quantized as the corresponding centroid of $\boldsymbol{c}_j$.

C. Principle of GMM-Based AVQ Scheme

Different from K-means clustering algorithm, which takes Euclidean distance as the decision condition for clustering and each sample only belongs to one cluster, the GMM divides the target samples into multiple different sub-Gaussian distribution models and gives the corresponding probability that each sample is assigned to M clusters. As a result, any sample belongs to multiple clusters at the same time, and finally the clustering decision is performed according to the maximum probability. The overall PDF of the GMM clustering established by the weighted sum of M Gaussian densities can be expressed as

$$P(\boldsymbol{x}|\boldsymbol{\lambda}) = \sum_{j=1}^{M} \omega_j \phi(\boldsymbol{x}|\boldsymbol{\mu}_j, \boldsymbol{\sigma}_j^2) \quad (4)$$
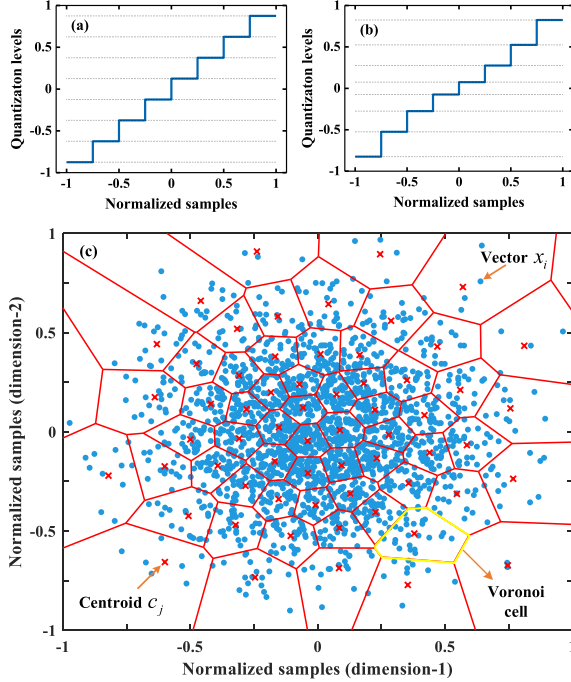
Fig. 1. Concept diagrams of (a) traditional uniform quantization, *K*-means-based (b) one-dimensional ASQ and (c) two-dimensional AVQ schemes.
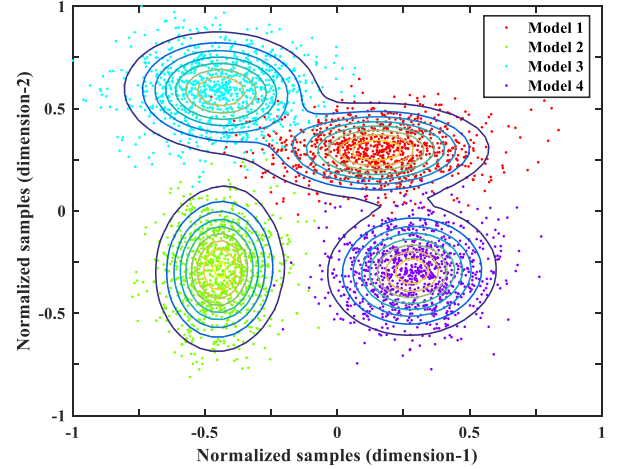

Fig. 2. Contour diagram of probability density function for two-dimensional GMM AVQ based on four sub-Gaussian models.

where ω_j is the mixture weight of the *j*-th component, and $\phi(\boldsymbol{x} \mid \boldsymbol{\mu}_j^2, \boldsymbol{\sigma}_j^2)$ is the Gaussian densities of *j*-th component, which meets the Gaussian distribution $N(\boldsymbol{\mu}_j, \boldsymbol{\sigma}_j^2)$ with the mean $\boldsymbol{\mu}_j$ and variance $\boldsymbol{\sigma}_j^2$. The mean and variance describe the distribution location and amplitude of the samples, respectively. The sum of all weights is equal to 1, i.e., $\sum_{j=1}^{M} \omega_j = 1$. It should be noted that the overall PDF estimation of GMM is given here just to help understand the principle of GMM, and it is not required in the actual GMM quantization process. The hidden parameters of the GMM are collectively represented by the notation of $\boldsymbol{\lambda}_j = \{\omega_j, \boldsymbol{\mu}_j, \boldsymbol{\sigma}_j^2\}, j = 1, 2, \ldots, M$. These parameters are preset with a randomly initial values, and then can be optimized by the expectation maximization (EM) algorithm [25], which contains two steps: step E and step M.

Firstly, we initialize the parameters of $\boldsymbol{\lambda}_j$ with $\omega_j = 1/M$, $\boldsymbol{\sigma}_j^2 = \boldsymbol{0}$, and $\boldsymbol{\mu}_j$ with the random quantization levels.

Secondly, using step E calculates the posterior probability γ_{ij} that the *i*-th sample $\boldsymbol{x}_i$ comes from the *j*-th sub-Gaussian model component, which can be expressed as

$$\gamma_{ij} = \frac{\omega_j \phi(\boldsymbol{x}_i | \boldsymbol{\mu}_j, \boldsymbol{\sigma}_j^2)}{\sum_{k=1}^{M} \omega_k \phi(\boldsymbol{x}_i | \boldsymbol{\mu}_k, \boldsymbol{\sigma}_k^2)}, \quad (1 \leq i \leq N, 1 \leq j \leq M) \tag{5}$$

Thirdly, using step M updates the parameter set of $\boldsymbol{\lambda}_j$ ($j = 1, 2, \ldots, M$) by

$$\omega_j = \frac{1}{N} \sum_{i=1}^{N} \gamma_{ij} \tag{6}$$

$$\boldsymbol{\mu}_j = \frac{\sum_{i}^{N} (\gamma_{ij} \boldsymbol{x}_i)}{\sum_{i}^{N} \gamma_{ij}} \tag{7}$$

$$\boldsymbol{\sigma}_j^2 = \frac{\sum_{i}^{N} \gamma_{ij} (\boldsymbol{x}_i - \boldsymbol{\mu}_j)(\boldsymbol{x}_i - \boldsymbol{\mu}_j)^T}{\sum_{i}^{N} \gamma_{ij}} \tag{8}$$

Next, repeat the step E and M process until the $\boldsymbol{\lambda}_j$ parameter change of two adjacent times reached the convergence threshold. Finally, the quantization process is performed according to the probability $\phi(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\sigma}^2)$, and each vector sample $\boldsymbol{x}_i$ can be matched to the corresponding mean μ of the sub-Gaussian model with the maximum probability. In theory, the GMM can fit any type of signal distribution without limiting the number of sub-Gaussian component. Fig. 2 shows the contour diagram of PDF for two-dimensional GMM-based AVQ, which enabled by four sub-Gaussian models. The target samples are generated through four kinds of Gaussian distribution with different colors. After GMM clustering, it can be seen that these samples can be divided into four sub-Gaussian models, and the contours of PDF for each sub-Gaussian distribution are in very high agreement with the distribution of the target samples, which means a good quantification effect can be achieved.

III. EXPERIMENTAL SETUP AND RESULTS

A. Experimental Setup

The experimental setup for 28 GHz photonics-assisted MMW communication system based on the low-bit non-uniform GMM AVQ scheme, is shown in Fig. 3. The intensity modulation with PAM4 signal and low-cost envelope detection are used to simplify the system structure. At the transmitting end, two external cavity lasers (ECLs) with the wavelengths of 1550.172 nm and 1550.396 nm are first coupled by an optical coupler (OC), and then are sent to an integrated Mach–Zehnder modulator (MZM) with built-in driving amplifier for electro-optic modulation. The PAM4 signal with the baud rate of 3 GBd is produced in the transmitting DSP and generated by an arbitrary waveform generator (AWG) with the sampling rate of 92 GSa/s. At the transmitting DSP, as shown in Fig. 3(a), the PAM4 symbols mapped from the pseudo-random binary sequence (PRBS) are
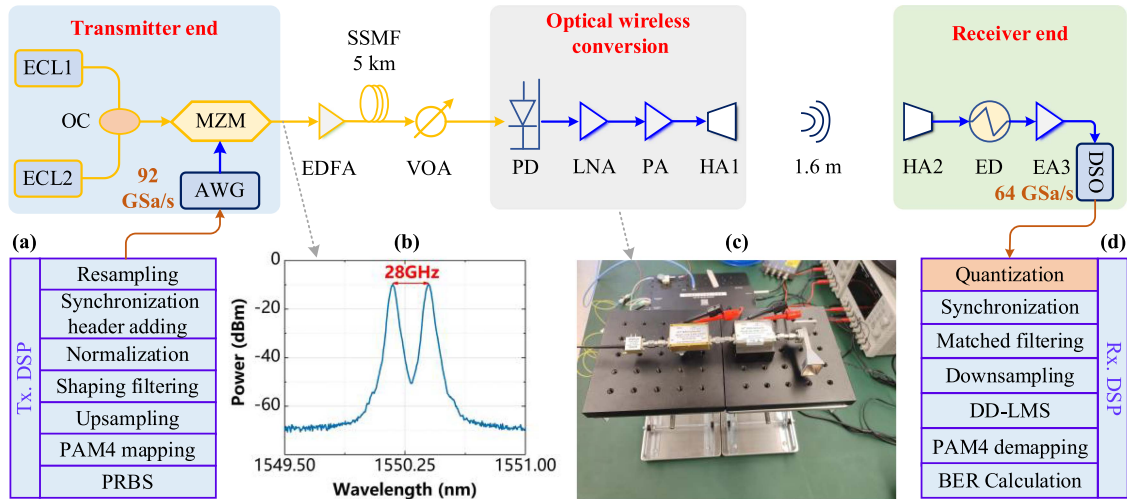
Fig. 3. Experimental setup for 28 GHz photonics-assisted MMW communication system based on low-bit non-uniform GMM AVQ technique. (a) Tx offline DSP. (b) Signal spectrum after electro-optic modulation. (c) Photo of optical wireless conversion module. (d) Rx offline DSP.
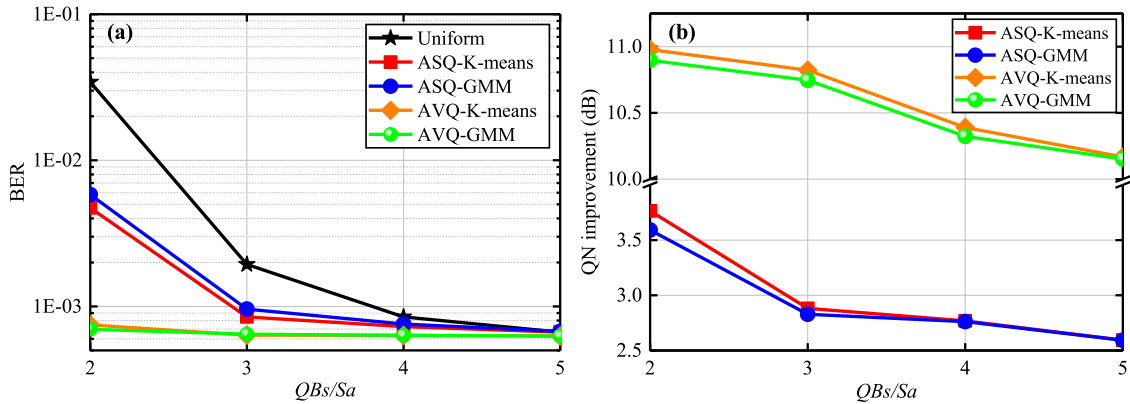

Fig. 4. (a) BER versus *QBs/Sa* curves for different quantization schemes. (b) Quantization noise (QN) improvement versus *QBs/Sa* curves for ASQ-*K*-means, ASQ-GMM, AVQ-*K*-means and AVQ-GMM schemes as compared with the uniform quantization scheme.

shaped by a root-raised-cosine (RRC) filter with a roll-off factor of 0.1 after 8 times up-sampling. This up-sampling operation guarantees the correlation between adjacent PAM4 samples, and thus can contribute to the performance improvement for multi-dimensional vector quantization. The modulated optical spectrum after MZM is shown in Fig. 3(b). The frequency interval between the two optical wavelengths is 28 GHz. Considering the photonics-assisted MMW indoor and outdoor continuous coverage scenario, a 5 km standard single-mode fiber (SSMF) is sufficient to provide the signal access from outside to indoor [26]. After fiber transmission, a variable optical attenuator (VOA) is applied to adjust the received optical power (ROP) for sensitivity measurement. At the optical wireless conversion side, whose photo is shown in Fig. 3(c), a single-ended PD with 3 dB bandwidth of 40 GHz is used to down-convert the optical signal into 28 GHz MMW through optical heterodyne beating. Afterwards, the generated 28 GHz MMW signal is launched into free space after amplified by one low-noise amplifier (LNA) and one power amplifier (PA) with the total gain of about 43 dB. A pair of horn antenna (HA) working in Ka-band (26.5∼40 GHz)

are employed for 1.6-m wireless transmission. At the wireless receiver end, the received 28 GHz MMW signal is reconverted to baseband via an envelope detector (ED) with the 3 dB bandwidth of over 500 MHz. Finally, the obtained baseband signal is amplified and then captured by a 64 GSa/s real-time digital storage oscilloscope (DSO) for further offline signal processing. At the receiving DSP, as shown in Fig. 3(d), the above signal is first quantized by a digital emulated ADC. For comparison, different quantization schemes including traditional uniform quantization, the present *K*-means and proposed GMM-based ASQ and AVQ are evaluated. Subsequently, the resampling, synchronization, matched filtering, downsampling, least mean square (LMS) equalization, PAM4 demapping and bit error rate (BER) calculation are performed in turn in the receiving DSP.

B. Experimental Results

We first investigate and compare the performance of the adaptive GMM quantization with other quantization schemes at different *QBs/Sa* condition under a fixed ROP. The results
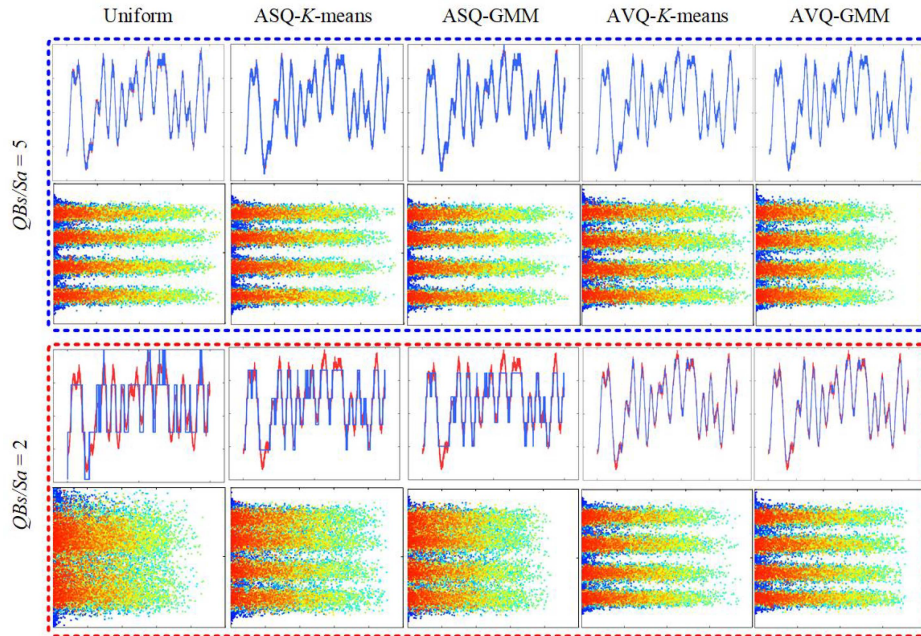
Fig. 5. The quantized time-domain waveform segments and recovered PAM4 level diagrams for different quantization schemes with the *QBs/Sa* of 5 bits and 2 bits, respectively.

are shown in Fig. 4(a). For the traditional uniform quantization scheme, the BER drops sharply with the increase of *QBs/Sa*, especially with an insufficient *QBs/Sa*. To be specific, the BER performance can be improved by more than an order of magnitude when the *QBs/Sa* is increased from 2 bits to 3 bits. Moreover, limited by the rigid quantization levels, the uniform quantization requires at least 5 bits to ensure a good and stable BER performance. On the contrary, benefitted from the flexible and adaptive quantization codebook, both the *K*-means and GMM-based ASQ/AVQ schemes can significantly reduce the requirement for *QBs/Sa*. For instance, the one-dimensional ASQ schemes can reduce the *QBs/Sa* requirement from 5 bits to nearly 3∼4 bits, whereas this value can be further reduced to 2 bits after utilizing the two-dimensional AVQ scheme. It can be found that the proposed GMM scheme can achieve the performance similar to *K*-means scheme, regardless of whether one-dimensional ASQ or two-dimensional AVQ technique is used.

In order to further investigate the quantitative performance of different quantization schemes, Fig. 4(b) shows the quantization noise (QN) improvement versus *QBs/Sa* curves for the *K*-means and GMM-based ASQ/AVQ schemes, as compared with the uniform quantization scheme. The QN improvement here is calculated according to the quantization variance, which represents the level of quantization distortion. That is, $\Delta QN = 10 \times \log_{10} (Var_0) - 10 \times \log_{10}(Var_x)$, where Var_0 and Var_x are the quantization variances of the uniform quantization scheme and non-uniform quantization scheme (including ASQ-*K*-means, ASQ-GMM, AVQ-*K*-means and AVQ-GMM), respectively. It can be seen from Fig. 4(b) that compared with uniform quantization, the QN improvement of ASQ and AVQ increases with the decrease of *QBs/Sa*. This is mainly because

the smaller the *QBs/Sa*, the larger the difference in quantization variance between rigid uniform quantization and ASQ/AVQ non-uniform quantization schemes. For one-dimensional ASQ scheme, the QN improvement is between 2.5 dB and 4 dB. Nevertheless, after adopting the two-dimensional AVQ scheme, the improvement can be increased to more than 10 dB. This implies better quantization performance can be achieved.

Fig. 5 shows the time-domain waveform segments and recovered PAM4 level diagrams after quantized processing by different schemes under a fixed *QBs/Sa*. In the time domain waveform segment, the processed waveform (blue curve) overlaps with the original waveform (red curve), thus the distortion level of the quantized sample can be observed intuitively. For a large *QBs/Sa* with 5 bits, the quantization distortion is insignificant even with the uniform quantization scheme. Meanwhile, there is also no significant difference in the level diagrams of five different quantization schemes. However, after reducing the *QBs/Sa* to 2 bits, the quantization distortions reflected by the waveform diagram for two-dimensional AVQ, one-dimensional ASQ and uniform quantization schemes increase successively. In particular, the performance of the uniform quantization at this case degrades notably, which can be observed from its cluttered level diagram. Instead, for the two-dimensional *K*-means and GMM-based AVQ schemes, although there is a slightly observable rise in quantization distortion, the influence on the level distribution is basically negligible, as compared with the case of 5 bits *QBs/Sa*.

We further evaluate the BER versus ROP curves under the case of *QBs/Sa* = 3 bits and *QBs/Sa* = 2 bits, respectively. The results are shown in Fig. 6. In order to clearly reveal the impact of quantization operation on the BER performance, the result without quantization is also given here as a reference. As
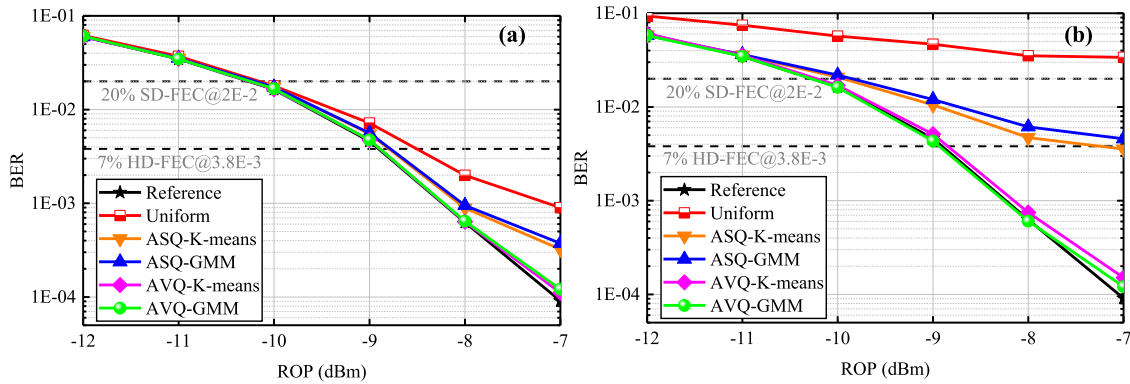
Fig. 6. BER versus ROP for without quantization (regarded as reference), uniform quantization, ASQ-*K*-means, ASQ-GMM, AVQ-*K*-means and AVQ-GMM schemes at (a) *QBs/Sa* = 3 bits, (b) *QBs/Sa* = 2 bits.

shown in Fig. 6(a) with the *QBs/Sa* of 3 bits, it can be seen that the BER performance of all schemes can meet the threshold of 7% overhead hard-decision forward error correction (HD-FEC) (3.8E-3). At this 7% HD-FEC BER threshold, the ASQ and AVQ schemes can improve the receiver sensitivity by about 0.3 dB and 0.4 dB, respectively, as compared with the performance of non-adaptive uniform quantization scheme. On the other hand, the BER performance of the proposed GMM-based ASQ and AVQ are basically similar to the *K*-means-based schemes. Meanwhile, compared with the reference result without quantization, the GMM and *K*-means-based AVQ have almost no performance penalty. When the *QBs/Sa* is reduced to 2 bits, it can be found from Fig. 6(b) that, the BER of non-adaptive uniform quantization scheme cannot meet the BER threshold of 20% overhead soft-decision forward error correction (SD-FEC) (2E-2). However, after utilizing the adaptive GMM and *K*-means-based ASQ schemes, the corresponding BERs can just reach the 7% HD-FEC threshold at the ROP of -7 dBm. Furthermore, the adaptive two-dimensional GMM and *K*-means-based AVQ schemes can further reduce the BER by more than an order of magnitude (below 2E-4) under the same ROP condition. At this case, the two AVQ schemes have also no performance penalty compared with the reference result, even if the *QBs/Sa* is as low as 2 bits. This means that the application of the two-dimensional AVQ scheme can significantly reduce the cost and power consumption of receivers widely deployed in optical wireless access networks.

Next, we evaluate the influence of different signal transmission baud rates on receiving performance for different quantization schemes. The results are shown in Fig. 7. The black curve is the reference result without quantization operation. We can observe the following three phenomena from Fig. 7. Firstly, under different baud rates from 1 GBd to 5 GBd, the proposed GMM quantization scheme can achieve comparable performance with the *K*-means scheme regardless of using one-dimensional ASQ or two-dimensional AVQ technique. Secondly, as the baud rate increases, the SNR performance of the two-dimensional AVQ scheme gradually and slightly deviates from the reference curve, especially when the baud rate exceeds 4 GBd. This is because the larger the signal baud rate, the higher the requirement for *QBs/Sa*. As a result, in the case of high baud rate, it is necessary to
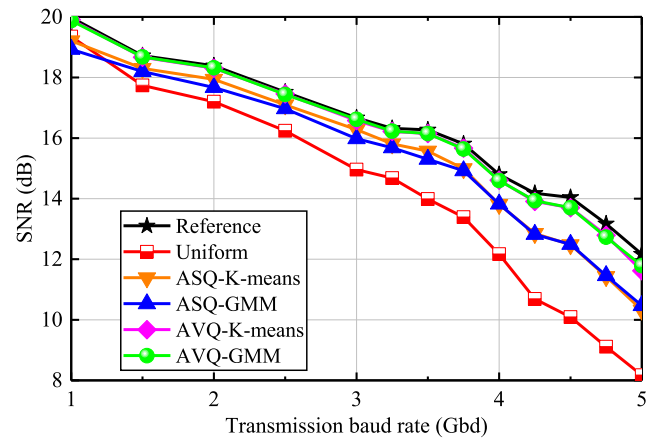

Fig. 7. SNR versus transmission baud rate for different quantization schemes. The black curve is the reference result without quantization operation.

increase the *QBs/Sa* value so as to ensure the same performance. Finally, with the increase of signal baud rate, the performance of the one-dimensional ASQ and two-dimensional AVQ are more obvious than that of the non-uniform quantization. In other words, the adaptive non-uniform quantization, especially multi-dimensional AVQ scheme, is more suitable for high-speed MMW communication scenarios.

Finally, we study the running time for *K*-means and GMM-based adaptive quantization schemes. In theory, since the GMM clustering algorithm utilizes the step E and M to estimate the posterior probability and sub-Gaussian model parameters, respectively, its complexity for a single iteration is higher than that of the *K*-means clustering algorithm. However, considering the case of a same convergence threshold, the elapsed time for the two kinds of clustering algorithms to reach the threshold is a more critical issue of our concern. As such, we test the running times of *K*-means and GMM-based one- or two-dimensional adaptive quantization under different *QBs/Sa* conditions. The results are shown in Fig. 8(a). It can be found that the quantization times of one-dimensional ASQ and two-dimensional AVQ schemes both rise with the increase of *QBs/Sa*. Meanwhile,
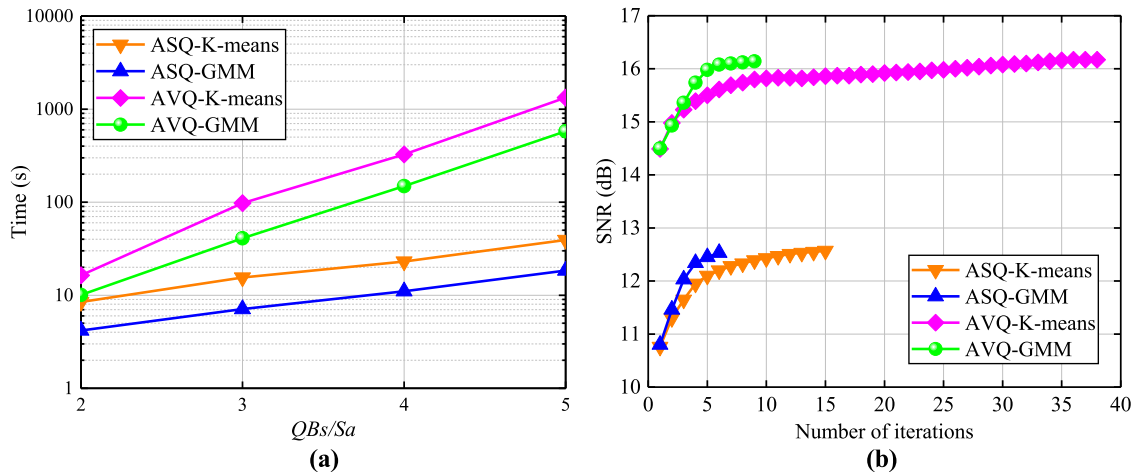
Fig. 8. (a) The running time for *K*-means and GMM-based adaptive quantization schemes. (b) The number of iterations for *K*-means and GMM-based adaptive quantization schemes to reach the same convergence threshold of 1e-3.

the bigger the *QBs/Sa*, the larger the time difference between two-dimensional AVQ and one-dimensional ASQ schemes. This main reason is that although two-dimensional quantization achieves the vector compression of scalar samples, it has more quantization levels than one-dimensional quantization in the same *QBs/Sa* case, thus the quantization time also increases significantly. On the other hand, the proposed GMM scheme has the shorter running time (nearly half) than *K*-means scheme regardless of using one-dimensional ASQ or two-dimensional AVQ technique. This might be mainly attributed to faster convergence of GMM clustering algorithm than that of *K*-means, which can be confirmed by Fig. 8(b). With the same convergence threshold of 1e-3, ASQ-*K*-means scheme requires 15 iterations to reach convergence, while only 6 iterations are enough for the ASQ-GMM scheme instead. For two-dimensional AVQ schemes, the iteration number of *K*-means and GMM-based schemes are 9 and 38, respectively, which also exists a multifold difference. The *K*-means adopts the hard decision method of Euclidean distance for clustering, while the distribution probability based soft decision is selected by the GMM approach. Relatively speaking, the decision for some samples in the edge of clusters using GMM algorithm is no longer rigid, but more flexible. This may promote faster convergence, especially for the signals that meet Gaussian distribution. In addition, it should be emphasized that, although there are differences in performance during the iteration process, the eventual SNR of the two schemes is basically similar after convergence.

IV. CONCLUSION

In conclusion, a photonics-assisted 28 GHz MMW communication system enabled by the GMM adaptive non-uniform quantization technique is proposed and verified experimentally. The GMM-based two-dimensional AVQ scheme is presented to realize the low-bit adaptive quantization of the ADC in the MMW receiver. We compare the performance of the proposed GMM quantization scheme with other quantization schemes (including the traditional non-adaptive uniform quantization,

the present *K*-means-based adaptive non-uniform quantization, etc.) in the low-cost intensity modulated envelope detection 28 GHz MMW communication system. The results show that, the quantization bit resolution of the ADC in our MMW receiver can be reduced as low as 2 bits by the GMM-based AVQ scheme, without noticeable performance penalty. In addition, the proposed GMM scheme can achieve a similar quantization performance as the *K*-means scheme. However, under the same convergence threshold, the quantization time can be reduced by about half. It should be emphasized that this proposed GMM-based multi-dimensional AVQ scheme is applicable to both pure electronics-enabled and photonics-assisted MMW communication systems. Nevertheless, relatively speaking, the effect of GMM AVQ scheme is more obvious in the latter case, which is characterized by high-speed and large-capacity scenario. Given the above, we believe it can provide an efficient and economic technical solution to reduce the cost and power consumption of ADCs in the upcoming B5G and 6G MMW-enabled optical fiber access networks.

REFERENCES

[1] X. You et al., "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Sci. China Inf. Sci.*, vol. 64, no. 1, Jan. 2021, Art. no. 110301.

[2] J. Yu, "Photonics-assisted millimeter-wave wireless communication," *IEEE J. Quantum Electron.*, vol. 53, no. 6, pp. 1–17, Dec. 2017.

[3] X. Li, J. Yu, and G.-K. Chang, "Photonics-aided millimeter-wave technologies for extreme mobile broadband communications in 5G," *J. Lightw. Technol.*, vol. 38, no. 2, pp. 366–378, Jan. 2020.

[4] P. Wang et al., "10 Gbps PAM4 transmission for fiber wireless access in Ka-band based on long short-term memory neural network equalizer," in *Proc. Int. Conf. Opt. Instruments Technol.*, 2022, pp. 88–94.

[5] M. Xu, Z. Jia, J. Wang, L. A. Campos, and G.-K. Chang, "Statistical data compression and differential coding for digital radio-over-fiber-based mobile fronthaul," *J. Opt. Commun. Netw.*, vol. 11, no. 1, pp. A60–A71, Jan. 2019.

[6] C. R. Berger, Y. Benlachtar, R. I. Killey, and P. A. Milder, "Theoretical and experimental evaluation of clipping and quantization noise for optical OFDM," *Opt. Exp.*, vol. 19, no. 18, pp. 17713–17728, Aug. 2011.

[7] N. S. André, H. Louchet, V. Filsinger, E. Hansen, and A. Richter, "Experimental 32 Gbaud OFDM transmission using a 3-bit DAC," in *Proc. Asia Commun. Photon. Conf.*, 2015, Paper no. ASu1E.1.

[8] B. Goebel, S. Hellerbrand, N. Haufe, and N. Hanik, "PAPR reduction techniques for coherent optical OFDM transmission," in *Proc. Int. Conf. Transp. Opt. Netw.*, 2009, pp. 1–4.

[9] W. Li et al., "32-GS/s 6-bit DAC based on SiGe technology for IM-DD OFDM systems with non-uniform quantization," *Opt. Exp.*, vol. 27, no. 6, pp. 8121–8129, Mar. 2019.

[10] D. Xue, W. Li, Z. He, M. Luo, H. Li, and X. Li, "Performance improvement of IM-DD OFDM systems by nonuniform DAC output levels," in *Proc. Asia Commun. Photon. Conf.*, 2018, pp. 1–3.

[11] P. Scheunders, "A genetic Lloyd-Max image quantization algorithm," *Pattern Recognit. Lett.*, vol. 17, no. 5, pp. 547–556, May 1996.

[12] M. Bi, T. Huang, L. Liu, X. Miao, S. Xiao, and W. Hu, "Performance optimization by nonparametric histogram estimation for low resolution in IMDD-OQAM-OFDM system," *IEEE Photon. J.*, vol. 10, no. 4, Aug. 2018, Art. no. 7905013.

[13] T. Huang et al., "Nonuniform quantification DAC for improving the performance of IMDD-based FBMC system," in *Proc. Int. Conf. Opt. Commun. Netw.*, 2017, pp. 1–3.

[14] J. Peng et al., "SQNR improvement enabled by nonuniform DAC output levels for IM-DD OFDM systems," *IEEE Photon. J.*, vol. 9, no. 2, Apr. 2017, Art. no. 7200411.

[15] Y. Yoffe and D. Sadot, "Novel low resolution ADC-DSP optimization based on non-uniform quantization and MLSE for data centers interconnects," *Opt. Exp.*, vol. 24, no. 5, pp. 5346–5354, Mar. 2016.

[16] H. Xu, M. Bi, F. Hu, Z. Hu, X. Teng, and W. Hu, "A GMM-based non-uniform quantization scheme for improving low-resolution IMDD-UFMC system performance," *Opt. Fiber Technol.*, vol. 71, May 2022, Art. no. 102943.

[17] H. Si, B. L. Ng, M. S. Rahman, and J. Zhang, "A novel and efficient vector quantization based CPRI compression algorithm," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7061–7071, Aug. 2017.

[18] W. Jiang, P. Liu, and F. Wen, "An improved vector quantization method using deep neural network," *Int. J. Electron. Commun.*, vol. 72, pp. 178–183, Feb. 2017.

[19] H. Ren et al., "Multidimensional vector quantization-based fast statistical estimation in compressed digitalized radio-over-fiber systems," *Appl. Opt.*, vol. 58, no. 13, pp. 3418–3425, May 2019.

[20] J. Ye, Y. Guo, P. Li, L. Yan, W. Pan, and X. Zou, "2-D quantization scheme utilizing SOFM neural network clustering for a DRoF system," *Opt. Lett.*, vol. 43, no. 19, pp. 4663–4666, Oct. 2018.

[21] L. Zhang et al., "Spectrally efficient digitized radio-over-fiber system with k-means clustering-based multidimensional quantization," *Opt. Lett.*, vol. 43, no. 7, pp. 1546–1549, Apr. 2018.

[22] L. Yue et al., "Improving performance of direct-detection terahertz communication system based on k-means adaptive vector quantization," in *Proc. Int. Conf. Opt. Commun. Netw.*, 2021, pp. 1–3.

[23] X. Feng et al., "Improving spectral efficiency of digital radio-over-fiber transmission using two-dimensional discrete cosine transform with vector quantization," *Opt. Exp.*, vol. 29, no. 16, pp. 25868–25875, Aug. 2021.

[24] J. Ye, Y. Guo, P. Li, L. Yan, W. Pan, and X. Zou, "Improving performance of digital mobile fronthaul employing 2-D vector quantization with vector linear prediction," *IEEE Photon. J.*, vol. 11, no. 2, Apr. 2019, Art. no. 7201911.

[25] G. Xuan, W. Zhang, and P. Chai, "EM algorithms of Gaussian mixture model and hidden Markov model," in *Proc. Int. Conf. Image Process.*, 2001, pp. 145–148.

[26] Y. Cai et al., "Demonstration of real-time photonics-assisted mm-wave communication based on Ka-band large-scale phased-array antenna and automatic beam tracking technique," in *Proc. Opt. Fiber Commun. Conf. Exhib.*, 2022, pp. 1–3.