ISSN: 2633-4828 Vol. 5 No.4, December, 2023

International Journal of Applied Engineering & Technology

ENHANCED AMPLITUDE PROBABILISTIC SHAPING BASED ON MODIFIED MULTI-REPEAT MAPPING FOR 5G MASSIVE-MIMO WIRELESS CHANNELS

Ali Shaban Hassooni[1, 2]* and Laith Ali Abdul Rahaim[1]
[1]Department of Electrical Engineering, College of Engineering, University of Babylon in Hilla, Iraq
[2]Department of Biomedical Engineering, College of Engineering, University of Babylon in Hilla, Iraq
*eng.ali.shaban@uobabylon.edu.iq

ABSTRACT
Probabilistic Amplitude Shaping (PAS) is one of the unique techniques that has attracted increasing interest in improving the performance of coded modulations in terms of improving spectral efficiency, reducing constellation energy, and bringing the achievable data rate closer to the Shannon limit. This paper will present a modified Probabilistic Amplitude Shaping (PAS) method based on a multi-repeat distribution matcher (MMRDM). This method will combine several small lookup tables to form a complete mapping table to change the probability of constellation points; this reduces the required storage capacity and computational complexity significantly, especially for High-level modulation schemes that include many energy levels such as 64-QAM (quadrature amplitude modulation) and above, this makes PAS technology suitable for the requirements of 5G and beyond; however, it is still not used in 5G systems, it is already used in optical communications and ATSC broadcasts. This paper will use the MMRDM with 16QAM and 64QAM according to the requirements and limitations of 5G NR standards with different levels, coding and shaping rates. We evaluate the method's performance regarding required storage capacity, complexity, and average bit error ratio (BER). As will seen, the simulation results of the proposed scheme show an improvement in the SNR of about (2-6.5 dB) at 1×10^{-6} BER, *for 16-QAM the entropy H={3,3.6}, for 64-QAM with entropy H={4.8,5.4} and various FEC rates compared with the uniform M-QAM and enumerative sphere shaping (ESS); this method offers a significant drop in the required capacity of storage memory and complexity of computation.*

Keywords: Probabilistic Shaping, MMRDM, QAM, 5G, Massive-MIMO.

I. INTRODUCTION
The significant increase in demands in terms of reliability, energy efficiency, capacity, spectrum efficiency, and other parameters has accelerated the deployment of the 5G wireless communication network [1]. This technological transfer to the new generation led to an exponential increase in productivity [2-4]. It has become urgent to mix and improve technologies to integrate with this generation and beyond to fulfil its requirements. Among these promising technologies is massive MIMO, which has recently attracted significant attention from researchers because of its ability to meet the previously mentioned requirements, overcome undesirable effects on the channel, and provide a reliable link without increasing the transmission power or bandwidth consumption. [3-5], so it can be considered without hesitation as one of the mainstays of fifth-generation systems [6].

In the new 5G radio network, the most problematic factor is to find the optimal modulation and coding schemes to accommodate the high data rate of the Gigabit range and higher, so the Quadrature Amplitude Modulation (QAM) is utilized in many ways depending on the users' distance from the central station [7-8]. However, the rectangular QAM signals can be easily generated from two PAM signals in the form of carriers in-phase and quadrature, in addition to the ease of demodulation. However, they have yet to consider the best M-ary QAM signals because the average power which provides the minimum distance is slightly greater than required in the optimal constellation of M-ary QAM. For this reason, the rectangular constellation QAM is still used even in 4G/LTE and 5G networks [9]. Using the uniformly distributed QAM constellation in the AWGN channel results in an energy efficiency reduction of up to 1.53 dB compared to Gaussian signals [10]. The gap in question is referred to as the "shaping gap" and denotes the disparity between the Shannon limit and the attainable data rate. Probabilistic shaping with a non-uniform distribution on the equidistant constellation points may be used to alleviate this issue. In this way, improving the bit error ratio (BER) and energy efficiency is possible, but there is a drawback: the

decrease in the bit rate. So, there is a trade-off between bit rate and power efficiency. The best distribution through which optimal energy efficiency can be obtained compared to the uniform distribution signal is the Maxwell-Boltzmann distribution (MB) [11]. Many researchers [12-15] proposed various transmission systems utilized probabilistic shaping, and the system performance has been improved in several respects, which confirms the preference for probabilistic shaping techniques in various communication systems.

Probabilistic amplitude shaping (PAS) coded modulation has emerged as a method of integrating modulation with channel coding. This approach has garnered significant attention in wireless and optical communications fields. The structure of probabilistic amplitude shaping, as described in [16], utilizes CCDM (Constant-Composition-Distribution Matching). The shaping technique used in this structure is arithmetic coding. The performance of CCDM increases with longer data packets [17].

Fehenberger et al. [18] presented an alternative concept to CCDM that is more appropriate when the block length is moderate, MPDM (Multipartition Distribution Matching), which is based mainly on the principle of CCDM. The MPDM matcher produces subsets of equal size by selecting compositions whose average is equal to the target compositions. MPDM increases the number of compositions that achieve the required distribution, so it performs better than CCDM in terms of rate losses and, thus, is better suited for applications with short block lengths. However, the compositions of numbers can be massive.

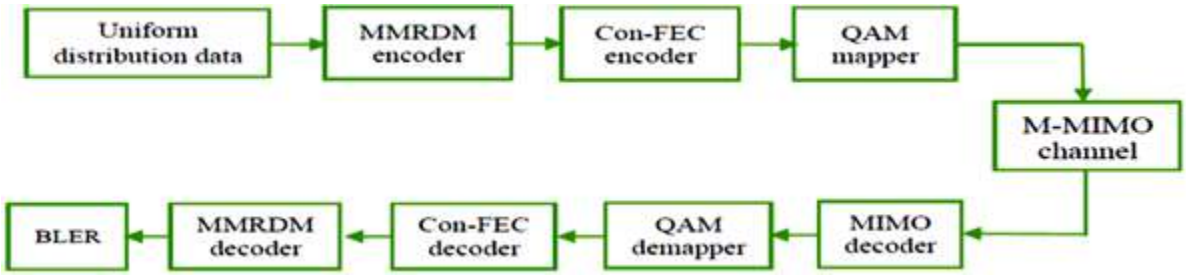


Figure 1. Transmitter and Receiver block diagram of probabilistic shaping 5G M-MIMO.

Based on this rate loss limitation, an improvement was made regarding rate loss at short to medium block length, where in [19], sphere mapping to match distribution (SMDM) was used by researchers to reduce rate losses at short data packets. In [14], a PAS-based enumerative sphere shaping (ESS) was proposed. It has been proven that the performance of ESS is better when compared to CCDM for small or medium block lengths when evaluated via the AWAGN channel, and the computational and storage complexity has also been reduced. However, some limitations and computational complexity remain, especially when long data packets.

Iscan et al. In [20], they used probabilistic shaping to improve system performance by introducing the shaping encoder before the polar encoder as a proposal for the 5G RAN (new radio network) based on the polar coding chains. This method uses shaping bits and attaching them to data bits to produce the targeted non-uniform distribution. The researchers used the polar decoder as a probabilistic shaping encoder to modify the existing block. On AWGN channels, More than 1dB improvements were obtained without increasing the required computing through simple operations.

In this paper, we will present probabilistic amplitude shaping based on the modified method, which is considered an improvement of the principle of multi-repeat distribution matching (MRDM) [21]. Several small lookup tables will be used to create a coding scheme, and in this process, the required storage capacity will be greatly reduced, and it does not need any complex calculations. To obtain the non-uniform distribution that follows the MB distribution, choose the appropriate number of levels, as well as the content of the subgroups, and the appropriate selection of the symbols that must be repeated at each level to achieve the required distribution so that a small lookup table is formed for each level so that a probabilistic encoder and decoder is formed according to the required rate of shaping. uniform distribution. The symbol distribution must be Gaussian-like to maximize the

Gaussian channel capacity. To study the performance of MMRDM, a mathematical analysis of the bit error ratio will be conducted through the AWGN channel, in addition to conducting simulations of the proposed method using MATLAB R2022b through the Massive-MIMO OFDM wireless channel with Convolution Coder at different FEC rates. The results show a significant reduction in storage capacity, computational complexity of the method, and error rate compared to uniform M-QAM and enumerative sphere shaping (ESS) for the same SNR (signal-to-noise ratio) and channel.

II. System Model

This paper adopts a coded modulation 5G Massive-MIMO communication system based on square QAM with PAS-MMRDM probabilistic shaping, as shown in Fig. (1). On the transmitter side, the PAS encoder reshapes the streamed bits with uniform distribution into the Gaussian-like distribution, As discussed in detail in the next section. Next, the convolutional forward error correction (conv-FEC) encoder processes the shaped stream, and the M-QAM mappers convert the bits into quadrature amplitude modulation (QAM) symbols.

After the signal passes through the Massive-MIMO channel and at the receiving end, the QAM demapper will demodulate the received signals based on the optimal maximum back-detector (MAP) because the transmitted signals are not uniformly distributed, and detection needs to take the priori probability of symbol into account, unlike a detector of Maximum likelihood that it does not do so. The FEC decoder receives the MAP signals from the demodulator to correct errors that occur in the channel, and then the original data bits are restored through the shape decoder.

III. Modified Distribution Matcher description

The principle of shaping scheme, as it is known, is to change the distribution of data from a uniform distribution to a non-uniform one that follows a specific distribution to improve energy efficiency by increasing the probability (repeating) of the constellation points with low energy and reducing the probability with those with higher energy. In [11], it was explained that if the probability of the constellation points follows the distribution in Eq. 1, better system performance will be achieved. This distribution, known as the Maxwell-Boltzmann (MB) distribution, is as below:

$$p(x) = A(\lambda)\, e^{-\lambda x^2} \quad x \in R \qquad (1)$$

Where λ is the rate parameter, trading off between data rate and average power, A(λ) is a parameter chosen such that $\sum p(x) = 1$.

The proposed method aims to achieve probabilistic shaping by adding extra bits to attain the one-to-one mapping technique, thereby reducing the computational complexity of the system. Instead of using sequence permutation, multiple repetition mapping of combinations is employed to acquire the mapping sequence in the PAS, known as multi-repeat shaping (MRS) [21]. The proposed modification in this paper is to populate multiple lookup tables with constellation points as subsets to generate a distribution that follows the MB distribution according to the target rate. Fig. (2) describes this modification illustratively.

To begin with, we determine the number of levels or groups (L) and then create sub-groups (S1, ..., SL) within those groups. Symbols consisting of bits are assigned according to the constellation's order M, representing the constellation's points. Lookup tables will be generated for each level or subset, and the probabilistic encoding process will done by combining the output of these tables from the original input data with a specific packet length [21].

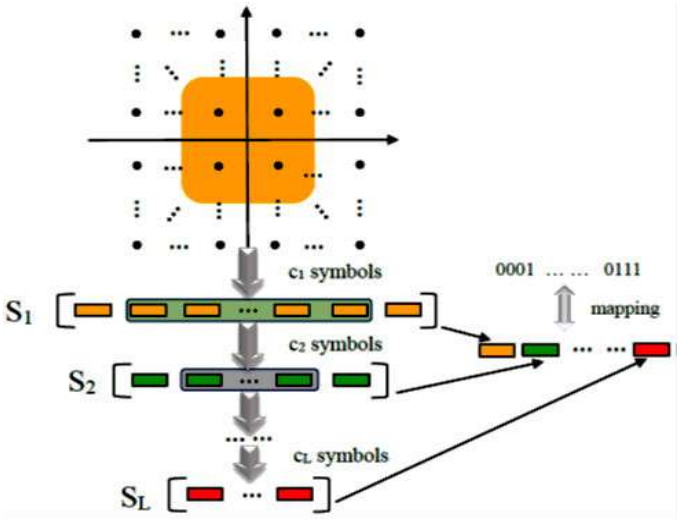


Figure 2. Shaping process.

When selecting symbols for constellations, it is essential to establish flexible rules in the shaping method. M symbols are used in an M-QAM constellation, each consisting of $N = \log_2(M)$ bits. Each level will contain a number of certain repeated symbols. For example, in the first level (S1), the number of repeated symbols will be c1. Then, the second subset (S2) will be formed from S1 and its number of symbols, c2, until the last group (SL) is reached with repeated symbols, cL. Following the energy rule when repeating symbols is essential, meaning symbols with lower energy should be repeated more frequently than higher energy symbols.

As previously mentioned, one-to-one mapping of symbols avoids the computational complexity in complex iterative methods when mapping many-to-one. Thus, symbols constructed from all subsets of a given block length will mapped with one-to-one mapping.

In order to achieve shaping encoding, this method requires the specification of several parameters [21]. The key parameters are L, which denotes the number of levels or groups; N, which represents the number of bits per symbol in the M constellation. The parameters (c1, c2, ..., cL), which signifies the number of symbols in each set group; and m, which stands for the block length of the input information.

Parameters are related to each other according to the following mathematical expressions:

$$\begin{cases} c_1 \le 2^N \\ c_2 \le c_1 \\ \dots\dots \\ c_L \le c_{L-1} \\ 2^m \le c_1 \times c_2 \times \dots \times c_L \le 2^{L \times N} \end{cases} \quad (2)$$

Lets $\zeta_{PA'}$ represents the set of vectors satisfied MB distribution PA'

$$\zeta_{PA'} = c_1 \times c_2 \times \dots \times c_L \quad (3)$$

Then m is the input length can be expressed as:

$$m = \lfloor log_2(c_1) \rfloor + \dots + \lfloor log_2(c_L) \rfloor \quad (4)$$

The formula of shaping rate (R) is.

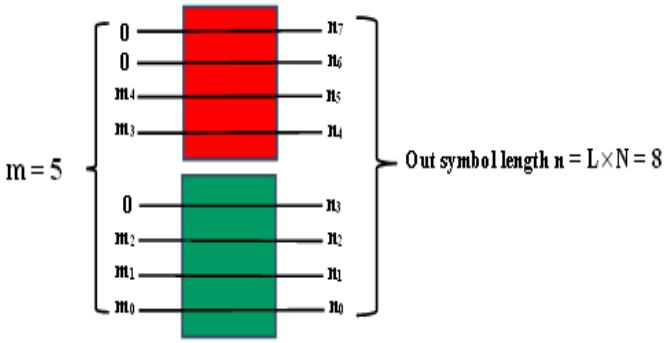
$$R = {}^{m}/_{L \times N} \quad (5)$$

The size of MMRDM codebook is $|C_{MRDM}| = 2^m$

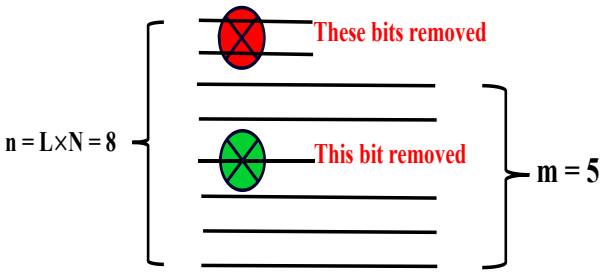
The parameter L governs the trade-off between the distribution divergence from the MB distribution on the one hand and the shaping encoder complexity on the other hand. So, the maximum length of combined codewords is L symbols, with N bits for each symbol. Therefore, the storage capacity $2^m \times (L \times N)$ bits is required for the created lookup table; this means that as the L levels increases and the modulation order increases, the computational complexity increases, and also the storage capacity will increase significantly, for example [21] if we set L to 2 and the modulation order is 16-QAM, where N=4, and levels c1=4, c2=8, according to Eq. (4), the length of the m block will be 5, so the lookup table size will be ($2^5 \times 8$) bits. This means the capacity required for storage will be small, but the distribution divergence will be greater. But if we set L to 3 for the same conditions and choose the groups as c1 = 4, c2 = 8, and c3 = 16, according to Eq. (4), the block length (m) become 9, the lookup table increases to ($2^9 \times 12$) bits, i.e., the storage capacity increases and the distribution divergence decreases. However, the required capacity will increase significantly if the modulation scheme increases. To clarify this fact, let us take 64-QAM, which contains ten different levels of energy and 64 constellation symbols; with N = 6, we will assume L = 7, and also the levels to be c1 = 4, c2 = 8, c3 = 16, c4 = 24, c5=32,c6=36,c7=64, using Eq. (4) it would be m=29 and ($2^{29} \times 42$) bits the lookup table size, which is impractical and its mathematical calculations are complicated.

We will simplify the lookup table by dividing it into smaller discrete tables. Each sub table will contain ci symbols, where i=1,2,..,L. For instance, when L = 2, the lookup table size is 256 bits [21, Table I]. However, with this modification, there will be two lookup tables; the first will contain 4 symbols, which represented by 4 bits each, while the next will contain 8 symbols denoted by 4 bits. This simplification means only 3 bits will be needed to store these encoding tables. Additionally, no table will be needed for the decoder, as illustrated in Fig. 3.

Similarly, when L=3, the lookup table size required will be 6144 bits to store, while only 3 bits will be needed if modified lookup tables are used to perform the coding. Fig 4 shows the shaping process in case L=3.

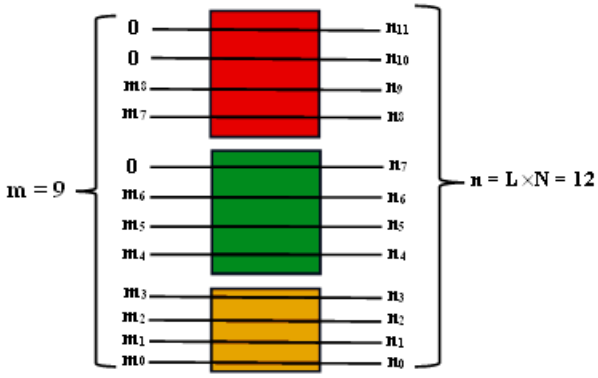


(a)

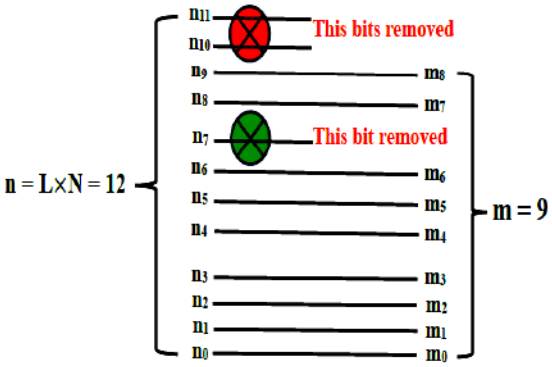


(b)

Figure 3: 16-QAM Shaping process with N=4, L=2, m=5, and c1=4, c2=8 a) encoding b) decoding.



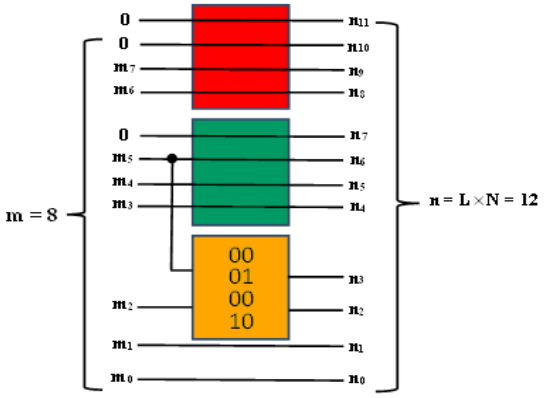
(a)



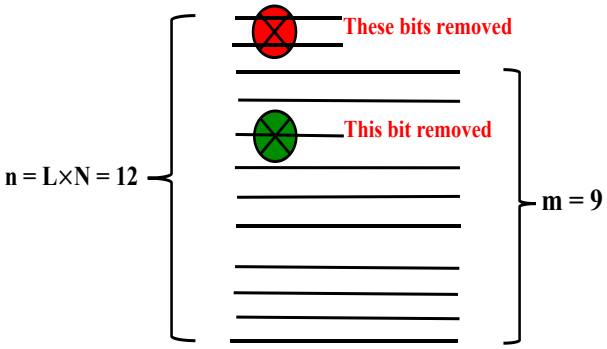
(b)

Figure 4: 16-QAM Shaping process with N=4, L=3, m=9, and c1=4, c2=8, c3=16 a) encoding b) decoding.

It is worth noting that the level or group that contains several symbols is not a power of 2. This case is considered a simple complication in calculating the search table for this group, where the lookup table for this level is chosen depending on the group with a power of 2 before it. Returning to the previous example, 16-QAM, where L = 3, but the third group will be c3 = 12, not power of 2, and packet length m = 8, where the lookup table size will be 3072 bits, while we only need 11 bits to represent it in the multiple coding lookup tables and 4 bit to decode lookup table, as shown in Figure 5. In the case of 64-QAM, L = 7, m = 29, the required saving capacity is 2.25485×10^{10} bits. In contrast, the required capacity will be 135 bits in the case of using multiple encoding and decoding lookup tables, and this is a significant decrease in the required memory, as illustrated in Fig. 6.

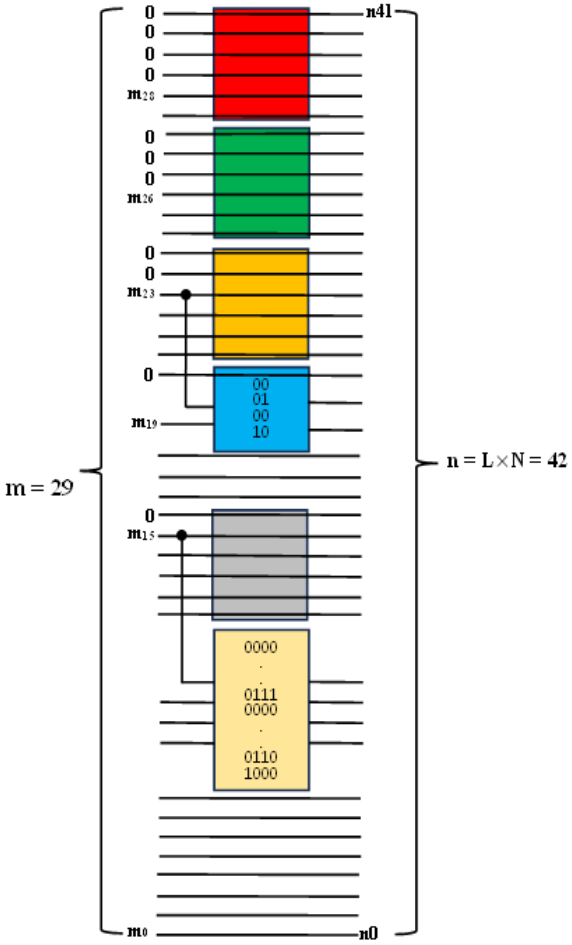


(a)

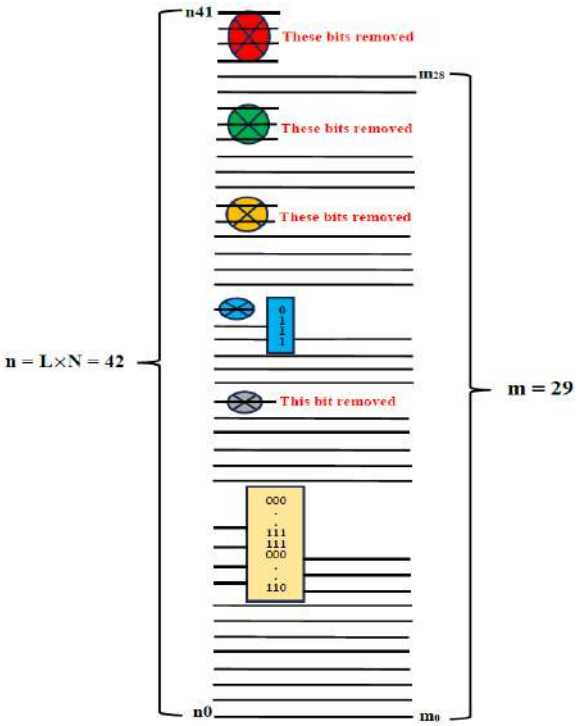


(b)

Figure 5: 16-QAM Shaping process with N=4, L=3, m=8, and c1=4, c2=8, c3=12 a) encoding b) decoding.



(a)



(b)

Figure 6: 64-QAM Shaping process with N=6, L=7, m=29, and c1=4, c2=8, c3=16, c4=24, c5=32, c6=36, c7= 64 a) encoding b) decoding.

Ultimately, the probabilities of the created distribution from the MMRDM may be computed as follows:

Shaped symbols probabilities are $p_A(a^M) = [p_{a1}, p_{a2}, \dots, p_{aM}]$, and there are L groups have different probabilities.

group g1 consists of c1 symbols with equal probabilities:

$$p_{g1}(a^{c1}) = [\frac{\frac{1}{c1}+\frac{1}{c2}+\dots+\frac{1}{cL}}{L}, \dots, \frac{\frac{1}{c1}+\frac{1}{c2}+\dots+\frac{1}{cL}}{L}]$$

group g2 consists of (c2-c1) symbols with equal probabilities.

$$p_{g2}(a^{c2}) = [\frac{\frac{1}{c2}+\dots+\frac{1}{cL}}{L}, \dots, \frac{\frac{1}{c2}+\dots+\frac{1}{cL}}{L}]$$

.

.

.

group g_L consists of (c_L-c_{L-1}) symbols with equal probabilities.

$$p_{gL}(a^{cL}) = [\frac{\frac{1}{cL}}{L}, \dots, \frac{\frac{1}{cL}}{L}]$$

Then the PS-QAM probabilities become

$p_A(a^M) = p_{g1} \cup p_{g2} \cup \dots \cup p_{gL}$.

For 16-QAM, L=2, m=8, c1=4, c2=8 the symbols probability will be

$$p_A(a^{16}) = [\frac{\frac{1}{c1}+\frac{1}{c2}}{L}, \frac{\frac{1}{c1}+\frac{1}{c2}}{L}, \frac{\frac{1}{c1}+\frac{1}{c2}}{L}, \frac{\frac{1}{c1}+\frac{1}{c2}}{L}, \frac{\frac{1}{c2}}{L}, \frac{\frac{1}{c2}}{L}, \frac{\frac{1}{c2}}{L}, \frac{\frac{1}{c2}}{L}, 0,0,0,0,0,0,0,0]$$

$$= [\frac{3}{16}, \frac{3}{16}, \frac{3}{16}, \frac{3}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, 0,0,0,0,0,0,0,0]$$

4. Proposed method performance

M-QAM's symbol error ratio (Ps) across the additive white gaussian noise (AWGN) channel, as it is known, can be represented as [22]:

$$P_s = 4\frac{\sqrt{M}-1}{\sqrt{M}} Q\left(\frac{d(M,E_b)}{\sqrt{2N_0}}\right) - \qquad 4\left(\frac{\sqrt{M}-1}{\sqrt{M}}\right)^2 Q^2\left(\frac{d(M,E_b)}{\sqrt{2N_0}}\right) \qquad (7)$$

Where d(M, Eb) represents the minimum distance that separates constellation symbols, it is determined by the number of symbols, M, and the bit energy, Eb. The number of bits per symbol, N, is equal to log2M. The symbols in the square M-QAM constellation are located at x = (±A, ±3A,…,±(√M-1)A) + i(±A, ±3A,…,±(√M-1)A). The average symbol energy for M-QAM is $Es_{QAM} = \frac{2}{3}(M-1)A^2$, whereas for MMRDM M-QAM, the average symbol energy is $Es_{MMRDM} = \sum_{i=1}^{M} p_{A_{MMRDM}}(a_i)\, E(a_i)$. Based on the above information, it can be inferred that Es_{MMRDM} is less than Es_{QAM}, and the minimum distance (dmin) equals 2A. Consequently, the value of A_QAM can be calculated using the formula $\sqrt{3Es_{QAM}/2(M-1)}$. This equation below represents the calculation of dmin.

$$dmin_{QAM} = 2\sqrt{3Es_{QAM}/2(M-1)}$$

$$= \sqrt{6Es_{QAM}/(M-1)} \qquad (8)$$

And $dmin_{MMRDM} > dmin_{QAM}$

Let's apply these conditions in 16-QAM as example.

$Es_{QAM} = 10A^2$

$dmin_{QAM} = \sqrt{2Es_{QAM}/5}$ sub. in Eq. (7)

16-QAM average symbol error rate

$$P_{s_{QAM}} = 3Q\left(\sqrt{\frac{Es_{QAM}}{5N_0}}\right) - \frac{9}{4}Q^2\left(\sqrt{\frac{Es_{QAM}}{5N_0}}\right)$$

Since $\boldsymbol{Es = \log_2(M)\, Eb = 4Eb}$

$$P_{s_{QAM}} = 3Q\left(\sqrt{\frac{4Eb}{5N_0}}\right) - \frac{9}{4}Q^2\left(\sqrt{\frac{4Eb}{5N_0}}\right) \qquad (9)$$

M-QAM Bit error ratio $Pb_{QAM} = \frac{P_{s_{QAM}}}{\log_2 M}$

For 16-QAM with MMRDM, L=2

$Es_{MMRDM} = 4A^2$, then the amplitude $A_{MMRDM} = \sqrt{Es_{MMRDM}/4}$

$dmin_{MMRDM} = \sqrt{Es_{MMRDM}}$ sub. in Eq. (7)

$$P_{s_{MMRDM}} = 3Q\left(\sqrt{\frac{2Eb}{N_0}}\right) - \frac{9}{4}Q^2\left(\sqrt{\frac{2Eb}{N_0}}\right) \quad (10)$$

Average Bit error ratio $Pb_{MMRDM} = \frac{P_{s_{MMRDM}}}{\log_2 M}$

By comparing equations 9 and 10, it is noted that the average symbol error ratio in MMRDM M-QAM is lower than in uniformly distributed M-QAM because the shaping process will minimize the average energy for the constellation symbols, and this leads to maximizing the minimum distance (dmin) between the constellation symbols compared to uniformly M-QAM.

In addition to the bit error rate, the performance of the proposed MMRDM probabilistic shaping is verified in terms of gain (asymptotically) using the figure of merit (CFM) of $\mathbb{C}$ constellation [11,23]:

$$G \triangleq \frac{CFM}{CFM_{\oplus}} = \frac{(2^{\beta}-1)d_{min}^2}{6E}, \beta \geq 2$$

$$= \left(1 - 2^{-\beta(p)}\right)\gamma_s(p)\gamma_c(\Lambda) \quad (11)$$

Where β represents the normalized bitrate, γ_s and γ_c represent the shaping and coding gain respectively. Since the geometry of the constellation itself did not change after shaping, then only γ_s (constellation shaping gain) is effective and is expressed as:

$$\gamma_s(p) = 2^{\beta(p)}V(\Lambda)^{2/N}/[6E(p)]$$

E(p) represents the mean energy per symbol, and $V(\Lambda)$ is the size of the area spread over by the points of the lattice with the probability distribution p(r). The effectiveness of the suggested approach may be assessed by computing the attainable shaping gain ratio (G_{MMRDM}), which is the ratio of the shaping gain produced by PS-MMRDM employing a QAM lattice of MMRDM size to the gain supplied by the MQAM size uniformly QAM constellation; it can be expressed as:

$$G_{MMRDM} = \frac{\gamma_{s_{MMRDM}}}{\gamma_{s_{QAM}}} = \frac{E_{QAM}}{E_{MMRDM}}$$

$$G_{MMRDM}dB = 10\log_{10}\left(\frac{\frac{1}{M_{QAM}}\sum_{i}^{M_{QAM}}|x_i|^2}{\sum_{j}^{M_{MMRDM}}|a_j|^2 p_{A_{MMRDM}}(a_j)}\right) \quad (12)$$

By substitution the previous values for $E_{QAM} = 10A^2$ and $E_{MMRDM} = 4A^2$, the gain ratio will be 3.9dB.

We observe two key factors when determining the necessary storage capacity for the suggested approach. If the chosen subsets are powers of 2, the needed storage capacity will be ($L \times \log_2(M-m)$). On the other hand, if any subset or group ci is not a power of 2, then the lookup table storage capacity required for this subset can be formulated as ($k_1 2^{k_2} + k_3$), where $k_1 = k_2 = \lfloor \log_2(ci) \rfloor - 1$ and $k_3 = \log_2(M - (\lfloor \log_2(ci) \rfloor + 1))$.

The storage capacity needed for ESS is directly proportional to the number of energy levels, denoted as L, that meet the shaping rate Rs. As L grows, the necessary storage capacity also increases. The necessary storage capacity may be represented as [14] $L(N+1)\lceil NR\_s \rceil$ bits.

Next is the computational complexity calculation, which represents the arithmetic operations the shaping methods need to perform encoding and decoding. For the suggested method, there are also two states. The first is if the levels or subgroups are for power 2, where we do not need any calculation process for the sender or receiver, as

illustrated in Figures 3 and 4. The second case is that the level is not to the power of 2. In that case, the operations required to determine the symbol's address in the lookup table are simple for both the transmitter and the receiver; then, the Time opacity for this subset is O(1). In comparison, the ESS requires $(|\mathcal{A}_m - \mathbf{1}|)\lceil NR_s \rceil$ bits operations per symbol for each transmitter or receiver, where $\mathcal{A}_m$ represents the cardinality.

According to the given information, choosing subgroups in the suggested approach with a power of 2 greatly decreases the amount of storage and computing complexity in comparison to ESS. For example, when there are 30 output symbols and the modulation method is 64-QAM with a modulation order of 120, the memory required for MMRDM is just 60 bits.

V. RESULTS AND DISCUSSION

In this proposed system, OFDM Massive-MIMO is considered for the base station by employing hybrid beamforming at the transmitter end of the communication system. Several techniques will be used for both single-user and multi-user systems. These techniques will sound the entire channel to provide CSI (channel state information) to the receiver by dividing the required pre-coding into two essential parts: the digital band and the analogue radio frequency components. The architecture of the 5G Massive-MIMO system consists of several essential parts: MMRDM Probabilistic shaping Encoder, Conv-FEC Encoder, QAM Mapper for channel estimation; LS Estimator will be used, Nearest Neighbor QAM Demapper, ZF Equalizer, shaping Decoder, to retrieve multiple data streams from the sender. Moreover, to find the system's performance by highlighting the BER, the shaping gain ratio and the normalized general mutual information are part of the parameters used to measure the reliability of the communication system.

In the MMRDM-5G Massive-MIMO system simulation, the channel model will be chosen with a carrier frequency of 28GHz and the data block size between 22464000 and 29,952,000 bits. Table 1 shows the parameters used in the MATLAB program.

Table 1. MATLAB Simulation parameters

Carrier frequency	**$28e^9$**
Delay spread	$1e^{-9}$
Cyclic prefix length	64
Num. guard carriers	[1:7 129 256-5:256]
FFT size	256
Num. user terminal antennas	64, 128
Num. base station antennas	64, 128
Num. OFDM symbols	16
QAM mapper	16, 64
FEC coding rate	1/2, 2/3, 5/6
Num. shaping levels	2,3 for 16-QAM 8, 10 for 64QAM

Figs. 7 and 8 show the relation between BER and SNR for 16QAM and 64QAM modulation, considering 64x64 transmitting and receiving antennas of the Massive-MIMO system. In Fig. 7, the MMRDM with sets or levels (L) is 2 and 3 and have entropies {3,3.6}, which are compared to the uniformly distributed 16-QAM, while in Fig. 8, the shaping groups (L) 8 and 10 with entropies {4.8,5.4} are adopted to be compared to the 64-QAM for three FEC rates. The number of antennas adopted in Figures 9 and 10 is 128 x 128 for the same probabilistic shaping conditions and coding rates. Probabilistic shaping based on MMRDM introduces higher performance than uniformly M-QAM in all systems, where the SNR improvement at 1×10^{-6} BER is about (2-6.5) dB for all MMRDM sets and various FEC rates.

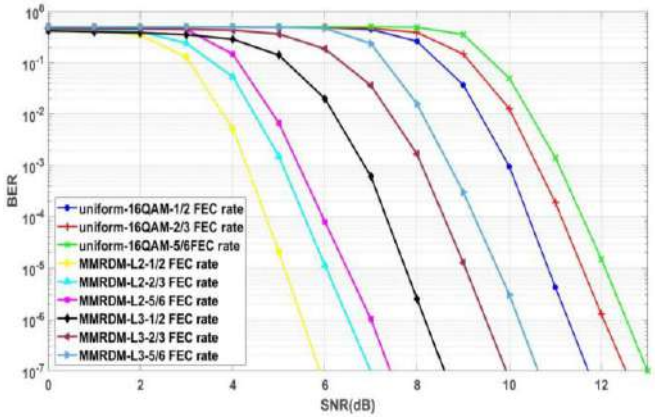


Figure 7. BER for M-MIMO systems with 64x64 antennas ,16-QAM, and MMRDM with L=2 and L=3.

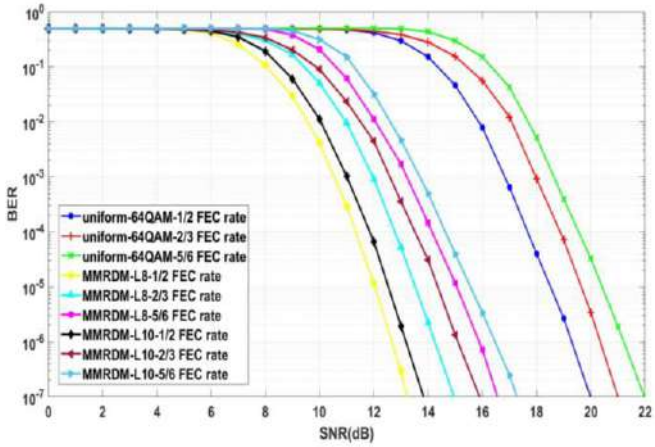


Figure 8. BER for M-MIMO systems with 64x64 antennas 64-QAM, and MMRDM groups L=8 and L=10.

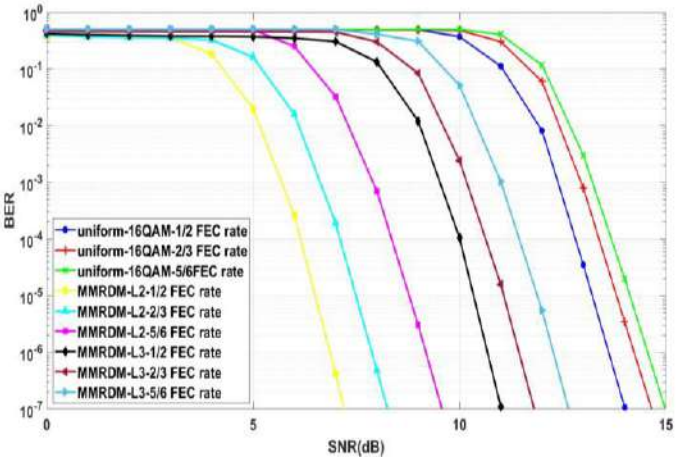


Fig. 9. BER for M-MIMO systems with 128x128 antennas ,16-QAM, and MMRDM groups L=2 and L=3.

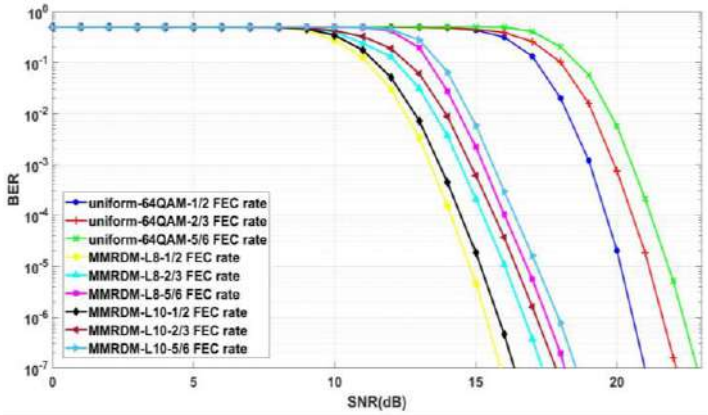


Figure 10. BER for M-MIMO systems with 128x128 antennas ,64-QAM, and MMRDM groups L=8 and L=10.

The comparison will now be made through the GMI values, where Figures 11and 12 plot the GMI vs SNR for the 16QAM and 64QAM mapping system. Noticeably, the value of the information exchanged in the MMRDM system is improved compared to the regular QAM. Except for MMRDM-16QAM, with given groups number L = 2, where the mutual information is less because the number of constellation points approved in the distribution is 8 out of 16, so the maximum GMI is 3 and not 4. Otherwise, an improvement in the GMI value is noted; for example, at 7dB, an improvement of 0.47 bit/symbol is observed. For MMRDM-L3 compared to regular 16-QAM, as in Fig. 11, the improvement is about 0.73 bit/symbol for MMRDM-L10 and 0.88 bit/symbol for MMRDM-L8 compared to 64-QAM for the same SNR value.

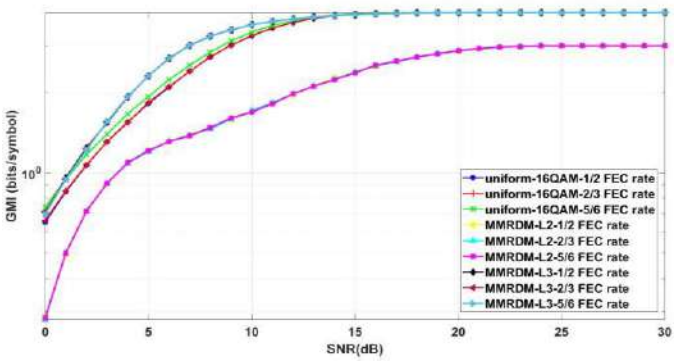


Figure 11. GMI vs SNR for M-MIMO systems with 64x64 antennas ,16-QAM, and MMRDM groups L=2 and L=3.

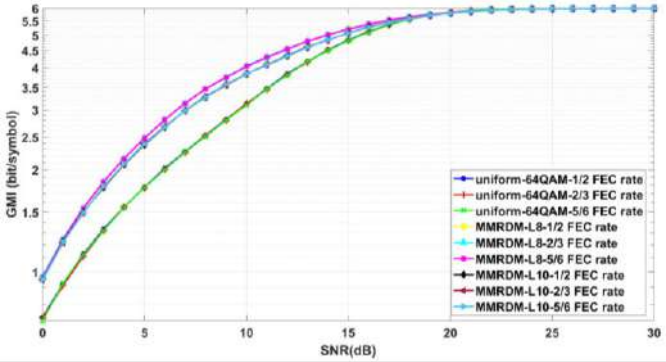


Figure 12. GMI vs SNR for M-MIMO systems with 64x64 antennas ,64-QAM, and MMRDM groups L=8 and L=10.

Table 2 shows the values obtained from comparing the results of the proposed method concerning the uniform M-QAM.

Table 2. The compared results of the proposed method to uniform M-QAM.

Mapping	MMRDM-16-QAM		MMRDM-64-QAM	
L	2	3	8	10
m=block length	5	9	34	44
n= NxL	8	12	48	60
rate =m/n	5/8	9/12	17/24	11/15
Storge capacity bits (encoding&decoding)	3	3	135	138
Gain ratio dB	3.9	2.28	4.1	3.8
Rate loss= H(Pa)-m/L	0.3096	0.2817	0.1392	0.1167

In the following, the BER performance of the proposed system will be evaluated at different block lengths with 16-QAM and 64-QAM modulation formats, taking into account the 64x64 transmitting and receiving antennas of the Massive-MIMO system. Figure 13 shows the performance of BER with a 16-QAM modulation scheme, 2/3 FEC coding rate, and symbol block length (4, 6, and 8). The proposed system improves BER compared to ESS and the system without modulation, where at 9dB, for example, MMRDM achieves 8×10^{-5} while ESS achieves 1.8×10^{-4}, which is much better than the system without modulation.

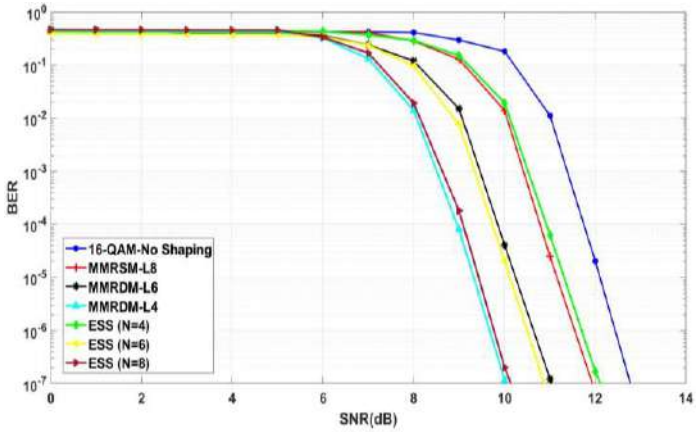


Figure 13. BER vs SNR for M-MIMO systems with 64x64 antennas ,and 2/3 FEC coding rate.

Figure 14 also shows the performance of the BER with a 64-QAM modulation scheme with a coding rate of 5/6 and the symbols block length (12,16,20). The proposed method also provides an excellent improvement compared to the system without shaping. The modified distributor achieves excellent performance with storage complexity much lower than ESS and low computational complexity on both the sending and receiving sides.

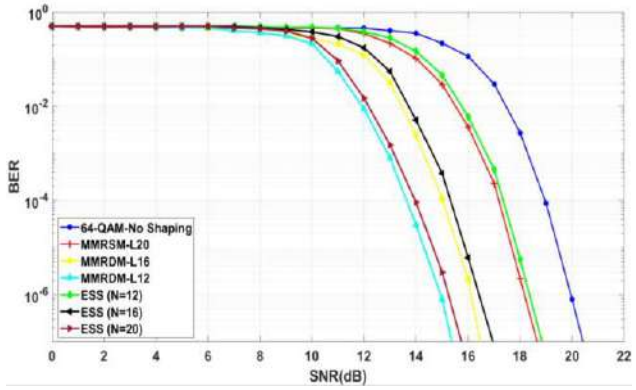


Figure 14. BER vs SNR for M-MIMO systems with 64x64 antennas, and 5/6 FEC coding rate.

VI. CONCLUSION

The probabilistic shaping method based on MMRDM was applied to the 5G Massive-MIMO communication system, where we discussed through the results the improvement in system performance, the simplified shaping process, and the lack of computational complexity or large storage capacity. However, it needs a small amount of memory compared to the length of the data block. Also, the energy efficiency is greatly improved when applying Eq. (11), which is the shaping gain ratio, in addition to a significant improvement in BER. The results and mathematical analysis show that the MMRDM method has good shaping advantages. However, it has one disadvantage: the rate loss is slightly high compared to ESS, so we need to optimize the values within the multiple lookup tables to reduce the rate loss; that is, we need to optimize the number of subsets or the L levels and the value of each subset within the lookup table and the method of its repetition.

REFERENCE

[1] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Commun. Surv. \& tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[2] N. Al-Falahy and O. Y. Alani, "Technologies for 5G networks: Challenges and opportunities," *It Prof.*, vol. 19, no. 1, pp. 12–20, 2017.

[3] [1] N. Lal, S. M. Tiwari, D. Khare, and M. Saxena, "Prospects for handling 5G network security: Challenges, recommendations and future directions," in *Journal of Physics: Conference Series*, 2021, p. 12052.

[4] [1] A. Sufyan, K. B. Khan, O. A. Khashan, T. Mir, and U. Mir, "From 5G to beyond 5G: A Comprehensive Survey of Wireless Network Evolution, Challenges, and Promising Technologies," *Electronics*, vol. 12, no. 10, p. 2200, 2023.

[5] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 742–758, 2014.

[6] M. Shafi *et al.*, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE J. Sel. areas Commun.*, vol. 35, no. 6, pp. 1201–1221, 2017.

[7] A. Al Amin, D. Basak, T. Khadem, M. D. Hossen, and M. S. Islam, "Analysis of modulation and coding scheme for 5th generation wireless communication system," in *2016 International Conference on Computing, Communication and Automation (ICCCA)*, 2016, pp. 1545–1549.

[8] E. Bobrov, D. Kropotov, H. Lu, and D. Zaev, "Massive MIMO adaptive modulation and coding using online deep learning algorithm," *IEEE Commun. Lett.*, vol. 26, no. 4, pp. 818–822, 2021.

[9] J. Wang *et al.*, "Capacity of 60 GHz wireless communications based on QAM," *J. Appl. Math.*, vol. 2014, 2014.

[10] H. M. de Oliveira, R. M. de Souza, and A. N. Kauffman, "Efficient multiplex for band-limited channels: Galois-Field division multiple access," *arXiv Prepr. arXiv1505.04140*, 2015.

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

[12] F. P. Guiomar *et al.*, "Adaptive probabilistic shaped modulation for high-capacity free-space optical links," *J. Light. Technol.*, vol. 38, no. 23, pp. 6529–6541, 2020.

[13] S. Civelli and M. Secondini, "Hierarchical distribution matching for probabilistic amplitude shaping," *Entropy*, vol. 22, no. 9, p. 958, 2020.

[14] Y. C. Gültekin, W. J. van Houtum, A. G. C. Koppelaar, F. M. J. Willems, J. Wim, and others, "Enumerative sphere shaping for wireless communications with short packets," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 2, pp. 1098–1112, 2020.

[15] E. Bobrov and A. Dordzhiev, "On Probabilistic QAM Shaping for 5G MIMO Wireless Channel with Realistic LDPC Codes," *arXiv Prepr. arXiv2303.02598*, 2023.

[16] 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, 2015.

[17] P. Schulte and G. Böcherer, "Constant composition distribution matching," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 430–434, 2015.

[18] T. Fehenberger, D. S. Millar, T. Koike-Akino, K. Kojima, and K. Parsons, "Multiset-partition distribution matching," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 1885–1893, 2018.

[19] P. Schulte and F. Steiner, "Divergence-optimal fixed-to-fixed length distribution matching with shell mapping," *IEEE Wirel. Commun. Lett.*, vol. 8, no. 2, pp. 620–623, 2019.

[20] O. \.I\cscan, R. Böhnke, and W. Xu, "Probabilistic shaping using 5G new radio polar codes," *IEEE Access*, vol. 7, pp. 22579–22587, 2019.

[21] Z. Jing *et al.*, "A multi-repeat mapping based probabilistic shaping coding method applied to data center optical networks," *Opt. Fiber Technol.*, vol. 61, p. 102401, 2021.

[22] K. Cho and D. Yoon, "On the general BER expression of one-and two-dimensional amplitude modulations," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1074–1080, 2002.

[23] G. D. Forney and L. F. Wei, "Multidimensional constellations-Part I: Introduction, figures of merit, and generalized cross constellations,'IEEE J. Select. Areas Commun., vol. 7, pp. 877-892, Aug. 1989. SI2 IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. 40, NO. 3, MARCH 1992 G. Ungerboeck," *A fast encoding method lattice codes quantizers," IEEE Trans. Inform. Theory*, vol. 29, pp. 820–824, 1983.