An Uplink SCMA Codebook Design Combining Probabilistic Shaping and Geometric Shaping

CHENGXIN JIANG, (Member, IEEE), AND YAFENG WANG, (Senior Member, IEEE)
Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Post and Telecommunications, Beijing 100876, China
Corresponding author: Yafeng Wang (wangyf@bupt.edu.cn)
This work was supported by the National Key Technology Research and Development Program of China under Grant 2018ZX03001022.

ABSTRACT Sparse Code Multiple Access (SCMA) is a promising non-orthogonal multiple access scheme owning to the shaping gain of its multi-dimensional codebook. Most of the existing designs of the codebook are based on geometric shaping (GS), while probabilistic shaping (PS) has important advantages in increasing channel capacity and reducing bit error rate. In this paper, an uplink SCMA codebook optimization algorithm which combines PS and GS is proposed. The algorithm adopts the bare bones particle swarm optimization algorithm based on maximizing the AMI. Based on the non-equiprobable codebook, a message passing algorithm with non-equalprobable distribution is introduced as the multiuser detection algorithm for SCMA scheme. We theoretically analyze the superiority of the codebooks combining PS and GS over others in terms of the average mutual information (AMI). Simulation results show that our proposed codebooks outperform the reference GS based codebooks under different overloading, different channels and different codebook sizes, and with the increase of overloading or codebook sizes, the gains are greater, which is up to 2.1dB gain in terms of block error rate performance. This observation also validates the theoretical analysis based on the AMI.

INDEX TERMS SCMA, codebook, probabilistic shaping (PS), geometric shaping (GS), average mutual information (AMI).

I. INTRODUCTION
Non-orthogonal multiple access (NOMA) schemes can effectively meet the stringent requirements of the fifth generation (5G) or beyond 5G wireless networks, which can improve the system capacity, user connectivity, and reduce service latency [1]. Sparse code multiple access (SCMA) is a promising multicarrier NOMA scheme, in which binary bits are directly encoded to multi-dimensional complex domain codewords of the predesigned codebook [2]. SCMA is an effective technique to improve spectral efficiency (SE), especially when used together with other techniques that improve SE, such as faster-than-Nyquist (FTN) [3]. Shaping gain of the multi-dimensional codebook is the main source of the performance improvement in comparison to the simple repetition of quadrature amplitude modulation (QAM) symbols in low density signature (LDS) [4], [5]. Due to the sparsity of the codebook, message passing algorithm (MPA) [4] can be applied to approximate the optimal SCMA decoder with much lower complexity.

As a means to improve SE, signal shaping has received extensive attention in recent years. It is generally divided into two categories: geometric shaping (GS) and probabilistic shaping (PS). In GS, the constellations are arranged in a non-evenly spaced way, and their probability distribution is uniform [6]–[8], while in PS, the constellations are arranged in an evenly spaced way, but their probability distribution is non-uniform [9]–[11].

Various research works have been done to improve the GS gain of SCMA by optimizing the codebooks. An SCMA codebook can be regarded as a multi-dimensional constellation. Unfortunately, the appropriate solution of this multi-dimensional problem is unknown, so designing an optimal SCMA codebook is still an open problem. A systematic sub-optimal method is used for SCMA codebook design [5]. In [12], based on maximizing the sum rate, a new design scheme of SCMA codebook is proposed to transfer the design of multi-dimensional constellation to the optimizing of series of one-dimensional complex codewords. Then the angles...
of rotation between the input one-dimensional constellation and the basic constellation are obtained with high computational complexity. A multi-dimensional SCMA codebook design based on constellation rotation and interleaving method was proposed for downlink SCMA systems [13]. Multi-dimensional constellation through signal space diversity (SSD) scheme was studied for uplink SCMA systems in [14]. Focusing on the optimized design of the mother codebook (MC), a novel codebook with large minimum Euclidean distance employing the star quadrature amplitude modulation (Star-QAM) modulation signal constellations is designed in [15], [16]. Joint optimization of constellation with mapping matrix is also proposed for SCMA codebook design [17].

The above literatures are all just based on GS to optimize the codebook, and the solution for a general and efficient multi-dimensional constellation construction has not been solved explicitly. For example, it is not clear how to obtain the rotation matrix for the lattice constellation in [5], while it is impossible to construct a multi-dimensional constellation of arbitrary size using the method in [14]. In [15], [16], the optimization criterion of codebook design is the large minimum Euclidean distance, which is not exact or considerable for NOMA schemes.

A PS based uplink NOMA scheme is proposed for fading channels to support multiple connections simultaneously [18], and a multi-step optimization scheme for maximizing the ergodic constellation constrained capacity is proposed to obtain the optimal probability mass function of these input signals. Compared with the uniform scheme, the simulations show that the proposed PS NOMA strategies can achieve great signal-to-noise ratio (SNR) gain in terms of the error performance. Inspired by this, we apply PS to SCMA codebook design, and to the best of our knowledge, little literature on this subject. In this paper, a joint codebook optimization algorithm combining GS and PS is proposed. This algorithm utilizes the bare bones particle swarm optimization (BBPSO) [19] algorithm based on maximizing the average mutual information (AMI). Then based on the non-equiprobable codebook, a Log-MPA multiuser detection algorithm with non-equiprobable distribution is introduced. Moreover, we theoretically prove the SE gains of the proposed codebooks over others through the AMI analysis, which can reflect the maximum information rate that can be transmitted without error under the given conditions [20].

The main contributions of this paper are summarized as follows:

- The AMI reflects the maximum information rate that can be transmitted without error under the given conditions, a joint BBPSO algorithm based on maximizing the AMI is proposed to optimize the geometric shaping parameters and the probability distribution of the SCMA codebook. In order to avoid the BBPSO algorithm falling into the local optima, the topology is modified.
- For different overloading, different fading channels and different codebook sizes, we fully prove the robustness of our proposed PS and GS based SCMA codebooks by AMI theoretical analysis and computer simulations, and with the increase of overloading or codebook sizes, the gains of the proposed codebooks over others are greater.

Throughout this paper we use boldface uppercase to represent matrices and boldface lowercase letters to represent vectors, and the scalar quantities are indicated by normal letters. The superscript \((\cdot)^T\) denotes transpose operation, and \(||\cdot||\) denotes the Euclidian distance. SNR \(= E_r/N_0\), let \(E_r\) denote the total energy per resource element at receiver, \(N_0 = 2\sigma^2\) denotes the one-sided noise power spectral density. \(x \in A^a\) denotes that \(x\) is a member of set \(A\) except for \(a\).

The paper is organized as follows. A system model of the PS and GS based uplink SCMA scheme and the Log-MPA multiuser detection algorithm with non-equiprobable distribution are introduced in section II. The detailed design process of our proposed SCMA codebook is discussed in section III. A brief introduction to other codebooks is also presented in this section. The AMI analysis and the joint BBPSO algorithm based on maximizing the AMI are presented in section IV. Simulation results are presented in section V. Section VI concludes the paper.

II. THE SYSTEM MODEL OF PS AND GS BASED UPLINK SCMA SCHEME

Fig.1 describes the structure of the PS and GS based uplink SCMA system. At the transmitter, the device constant composition distribution matcher (CCDM) is utilized to convert the uniformly distributed sequence of data bits into a sequence of non-uniformly distributed symbols, for detailed description please refer to [21]. Then an SCMA encoder for user \(j\) directly maps the encode bits to the codeword \(x_j\) of the non-equiprobable codebook,(which will be described in more detail later). After passing through the uplink single-input single-output (SISO) fading channel, at receiver we assume each user only has one data layer, and symbol level synchronization is assumed, thus the received signal can be expressed as:

\[
y = \sum_{j=1}^{J} \text{diag}(h_j)x_j + n, \tag{1}
\]

where \(x_j = [x_j^1, \ldots, x_j^K]^T\) is the SCMA codeword of user \(j\), \(h_j = [h_j^1, \ldots, h_j^K]^T\) is the corresponding channel vector of user \(j\), \(K\) is the number of subcarrier and \(n \sim CN(0, N_0 I)\)
is the additive complex zero-mean Gaussian noise with single-side power spectral density $N_0$.

To make full use of the prior information brought by the non-equiprobable codebook, based on the original MPA introduced in [4], we give a Log-MPA multiuser detection algorithm with non-equiprobable distribution here. As described in [4], for higher order modulation MPA, the message being exchanged must be in the form of a vector of size $M (M - 1)$ when taking one value as reference to others, and $M$ is the size of the codebook comprising the reliability values for each of the possible values taken from the codebook $CB$. Let $\xi_k$ and $\varsigma_j$ be the set of 1’s position in the $k$-th row and in the $j$-th column of the factor graph matrix, respectively. Let $L_{c_k \rightarrow u_j}^m$ and $L_{v_{k} \leftarrow u_j}^m$ be the $m$-th ($m = 1, \ldots, M - 1$) message sent from variable node $u_j$ and function node $c_k$, respectively. The message $L_{v_{k} \leftarrow u_j}^m$ gives an updated inference of $x_k^j$ based on the observation taken at chips $y_t$, $\forall t \in \varsigma_j \setminus \xi_k$.

Then $L_{v_{k} \leftarrow u_j}^m$ can be written as:

$$L_{v_{k} \leftarrow u_j}^m = \sum_{i \in \varsigma_j \setminus \xi_k} L_{v_{k} \leftarrow u_j}^m,$$ (2)

At the function node $c_k$, the message $L_{v_{k} \leftarrow u_j}^m$ can be written in (3), in the following page, where $x_k^i = [x_{j_1}^k, \ldots, x_{j_d}^k]^T$, $j_i \in \xi_k$, $i = 1, \ldots, d_f$ and $h_i^k = [h_{j_1}^k, \ldots, h_{j_d}^k]^T$, $j_i \in \xi_k$, $i = 1, \ldots, d_f$ represent the vector of symbols participating in the $k$-th resource element and their corresponding channel coefficient vector respectively. $d_f$ denotes the maximum number of users that is allowed to interfere within a single chip. $CB_j$ denotes the symbol in the $k$-th subcarrier and the $m$-th mapping symbol of the $j$-th user’s codebook $CB_j$, $p(x_i^k, x_j^k = CB_{j}^{km})$ represents the probability of the superposition vector is $x_i^k$ and $x_j^k = CB_{j}^{km}$, which can be detailed written as:

$$p(x_i^k, x_j^k = CB_{j}^{km}) = p(x_i^k = CB_{j}^{km}) \prod_{i \in \xi_k \setminus \varsigma_j} p(x_i^k),$$ (4)

where the probability $p(x_j^k = CB_{j}^{km})$ is equal to the probability of the $m$-th mapping vector of the non-equiprobable codebook $CB_j$. Equation (4) is a modification to the Log-MPA algorithm based on the non-equiprobable codebook, which is not necessary for the equiprobable one. It is a special case in the absence of feedback external information.

III. SCMA CODEBOOK DESIGN PROCESS

The performance of the codebook is crucial to SCMA scheme. Shaping gain of the multi-dimensional codebook is the main source of the performance improvement in comparison to the simple repetition of QAM symbols in LDS [4], [5]. Next, we will give the detailed construction process of our proposed non-equiprobable codebooks, and before that, as the compared schemes, we will give a brief introduction to the codebooks design in [5] and [16].

A. CODEBOOK OF ORIGINAL SCMA [5]

The design of SCMA codebook is a multi-dimensional problem and the appropriate solution of which is unknown now. A multi-stage optimization approach is proposed to achieve a sub-optimal solution for the problem in [5]. It includes mapping matrix design based on the factor graph matrix, multi-dimensional mother constellation design and constellation function operators.

The design of codebook is based on the factor graph matrix, two examples of typical and popular factor graph matrices with overloading factor $\eta = J/K = 150\%$ and $\eta = 200\%$ used
in [5], [16] are:

\[
F_1 = \begin{bmatrix}
1 & 1 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 1
\end{bmatrix}, \quad F_2 = \begin{bmatrix}
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1
\end{bmatrix}.
\]

(5)

(6)

When the factor graph matrix \( F \) is given, the unique solution of mapping matrix \( V_j \) is simply determined by inserting \( K - N \) all-zero row vectors within the rows of \( I_N \). \( N \) is the number of non-zero entries in the codeword. After getting the mapping matrix \( V_j \), the optimization of the codebook is reduced to a mother constellation and layer-specific operators.

As described in [5], the design objective includes both the sum distance and the product distance between the points by unitary rotation in the mother constellation. For higher rates, a structured construction approach is required to design the basic constellation. Then a shuffling method is proposed to construct an \( N \)-dimensional complex mother constellation from Cartesian product of two \( N \)-dimensional real constellations.

By having a solution for the mother constellation, the SCMA optimization problem is further reduced to optimize the layer specific operator \( A_l \). In uplink, as the layers pass through different channels, layer specific operator loses its importance [5], so we won’t reiterate it here for detailed construction process, please refer to [5].

### B. CODEBOOK BASED ON STAR-QAM SIGNALING CONSTELLATIONS [16]

The authors of [16] design the mother codebook of Star-QAM signaling constellation (SQ-SCMA codebook) based on the Star-QAM constellation, whose aim is to obtain a combination constellation with large minimum Euclidean distance. The steps to design SQ-SCMA mother codebook are roughly illustrated as follows:

**Step 1:** Design four vectors as

\[
t_1 = \left[ r_1, r_2, \ldots, r_M \right]^T, \quad t_2 = \left[ r_M + 1, r_M + 2, \ldots, r_M \right]^T,
\]

\[
i_1 = \left[ r_M - 1, \ldots, r_1 \right]^T, \quad i_2 = \left[ r_M - 1, \ldots, r_M + 1 \right]^T.
\]

(7)

where

\[
r_l = ((l - 1)(\alpha - 1) + 1), \quad l = 1, 2, \ldots, \frac{M}{2},
\]

and \( \alpha \in (1, +\infty) \) is a parameter that increases the amplitude linearly, which is detailed design in [16].

**Step 2:** Then the \( t \)-th dimension of SQ-SCMA codebook can be defined as follows, \( t = 1, 2, \ldots, N \) :

\[
g_t = \begin{cases}
\bar{t}_2^T - t_1^T & , \quad t \text{ is odd} \\
-\bar{t}_2^T & , \quad t \text{ is even}
\end{cases}
\]

(9)

**Step 3:** At last, the mother codebook, \( A_{MC} \), can be written as:

\[
A_{MC}^{N \times M} = e^{i\theta} [g_1, g_2, \ldots, g_{N-1}, g_N]^T,
\]

(10)

where \( \theta \in [0, 2\pi] \), and \( \mu \in (1, +\infty) \) is a parameter jointly designed with \( \alpha \) in [16] to get large minimum Euclidean distance of the combination constellation. Fig. 2 shows an example of SQ-SCMA codebook with \( N = 2 \)-dimensional and size \( M = 4 \), where the radii are given by \( r_1, r_2, r_1', r_2' \) respectively, and \( r_1' = \alpha r_1, r_2' = \alpha r_2, r_2 = \mu r_1 \). As described in [16], when \( \alpha = 3, \mu = 1.6 \) the performance of the codebook is best. Thus, the basic codebook, \( A_{MC} \), can be expressed as

\[
A_{MC} = \begin{bmatrix}
\alpha r_1 & r_1 & -r_1 & -\alpha r_1 \\
-r_2 & \alpha r_2 & -\alpha r_2 & r_2
\end{bmatrix}.
\]

(11)

![FIGURE 2. Illustration of the mapping function of MC SQ-SCMA codebook.](image-url)

After obtaining the MC, for uplink scheme, the design of the SCMA code book is left with a layer-mapping operator.
i.e., inserting \(K - N\) all-zero row vectors into \(A_{MC}\) according to the factor graph matrix \(F\).

C. CONSTRUCTION PROCESS OF OUR PROPOSED CODEBOOK

The two codebook design methods introduced above are only based on GS. However PS has important advantages in increasing channel capacity and reducing bit error rate. Therefore, a codebook design method combining PS and GS will be introduced next. As we all know, in SCMA scheme, binary bits are directly encoded to multi-dimensional complex domain codewords of the predesigned codebook. To construct a multi-dimensional SCMA codebook with codebook size \(M\), it needs a vector \(s\) with \(\log_2 M\) bits to finish the mapping from bits to a codeword. Let’s divide the \(\log_2 M\) bits into \(Z\) parts, then \(s = [s_1, s_2, \ldots, s_Z]\). The 1-dimensional \(M\)-points basic constellation \(c_1\) in the codebook can be obtained by the Cartesian Product of \(Z\sqrt{M}\)-order pulse-amplitude modulation (\(\sqrt{M}\)-PAM) constellation sets, and the \(\sqrt{M}\)-PAM constellation \(cp\) can be written as:

\[
cp = \{cp_1, cp_2, \ldots, cp_{\sqrt{M}}\} = \{-r \sqrt{M}, \ldots, -r, -r_1, r_1, r_2, \ldots, r \sqrt{M}\},
\]

and

\[
a_m = |r_{m+1}|/|r_1|, \quad m = 1, 2, \ldots, \sqrt{M}/2 - 1,
\]

where \(a_m \in (1, +\infty)\) is the parameter of amplitude increase, which is very important for SCMA codebook design. In [16], there is only one \(\alpha\) parameter, and when \(\alpha\) is designed, the amplitude sets of \(r_1\) in (7) is a simple arithmetic sequence, while \(a_m\) in our proposed codebook are different and jointly optimized with each other to get a more efficient amplitude sets. This paper proposes a joint optimization algorithm to obtain the optimal values of \(a_m, \mu_i\) (to be introduced later) and the probability distribution based on maximizing the AMI. The detailed algorithm is introduced in Algorithm 1.

For each quadrant-symmetric basic constellation, it has \(Q = 2^Z\) quadrants and \(M/Q\) points in each quadrant. The element \(s\), mapped to the PAM constellation is labeled under the gray rule with \(\hat{i} = (\log_2 M)/Z\) bits, i.e., \(s_z = [s_{1z}, s_{2z}, \ldots, s_{\sqrt{M}z}], \ z = 1, \ldots, Z\). Then the mapping rule of \(s\) is:

\[
s = [s_1, s_2, \ldots, s_Z] \rightarrow [b_Q, b_G]
\]

\[
= [s_1^1, s_2^1, \ldots, s_\sqrt{M}^1, s_1^2, s_2^2, \ldots, s_\sqrt{M}^2, \ldots, s_1^Z, s_2^Z, \ldots, s_\sqrt{M}^Z].
\]

Thus, the vector \(b_Q\) with the first \(Z\) bits represents the quadrant mapping. Vector \(b_G\) with \(\log_2 M - Z\) bits represents the relative location or amplitude of the constellation points in the quadrant. According to the quadrant symmetric property, the probability of the codeword mapped by \(s\) can be written as:

\[
p(s) = p(b_Q)p(b_G).
\]

where the probabilities of the quadrants are equiprobable, i.e., \(p(b_Q) = 1/Q\), and \(p(b_G) = p_{a_u}\) with \(b_G\) mapping to \(a_u\), where \(p_{a_u}\) is the probability of the non-equiprobable symbols with amplitude value \(a_u\), which will be optimized in Algorithm 1. Based on (12) to (15), we can construct the non-equiprobable 1-dimensional constellation \(c_1\) of the codebook.

After completing the design of the 1-dimensional constellation \(c_1\), scaling is used to construct other dimensions of the codebook, which can ensure different non-zero dimensions of codebook are not simple repetition, and therefore shaping gain can be obtained. Besides, power imbalance across the dimensions of codebooks helps MPA detector to operate more effectively to remove interferences, therefore the parameters \(\mu_i, i = 1, \ldots, N - 1\) are designed to make the energy of each dimension of the mother codebook different with each other, then the mother codebook can be written as:

\[
C = a[c_1, \mu_1c_1, \ldots, \mu_{N-1}c_1]^T,
\]

where \(a\) is a predetermined constant decided by the average energy and \(\mu_i \in (1, +\infty)\) is a parameter which will be jointly optimized with \(a_m\) and the probability distribution to get the maximum AMI.

For the PS and GS based SCMA scheme, firstly, user \(j\) utilize device CCDM to transform \(q\) uniformly distributed data

---

**Algorithm 1** The Joint BBPSO Algorithm of PS and GS Based Codebook by Maximizing the AMI

**Require:** \(W_s\): the size of the swarm; \(D\): the dimensions of the search space, i.e., the number of optimization parameters; \(M_I\): the maximum number of iterations.

1: \(t = 0\), randomly initialize \(q_{ij} \in (1, +\infty), j \in [1, Z(\sqrt{M}/2 - 1)],\) and \(q_{ij} \in (0, 1), j \in [Z(\sqrt{M}/2 - 1) + 1, D]\), \(P_i = Q_i, i \in [1, \ldots, W_s]\), and \(Q_i = [q_{i1}, \ldots, q_{iD}]\) is a \(D\)-dimensional set of all the \(D\) coordinates for the \(i\)-th particle. \(T_i = \arg \max(f(P_i(t))), f(\cdot)\) is the AMI \(P_{V_{i, e_i L_i}}\) computational formula (22), which depends a lot on \(q_{ij}\), and \(q_{ij}\) is a geometric shaping parameter or a probability of the codeword in the proposed codebook.

2: for \(t = 1\) to \(M_I\) do

3: \(i = 1: W_s\) do

4: Update the position of the \(i\)-th particle using (25);

5: end for

6: for \(i = 1: W_s\) do

7: \(P_i(t + 1) = \begin{cases} Q_i(t + 1), & f(Q_i(t + 1)) > f(P_i(t)) \\ P_i(t), & \text{else} \end{cases}\)

8: \(T_i(t + 1) = \begin{cases} P_i(t + 1), & f(P_i(t + 1)) > f(T_i(t)) \\ T_i(t), & \text{else} \end{cases}\)

9: \(G(t + 1) = \arg \max(f(T_i(t + 1)))\)

10: end for

11: end for

12: return \(G\) i.e., the set of parameters for geometric shaping and the optimal probability distribution, as well as \(f(G)\) i.e., the maximal AMI;
bits $b_j^n = \{b_j^{m+1}, \ldots, b_j^{m+q}\}$ into $n$ amplitudes (or location of the symbols) $\{A_j^1, \ldots, A_j^n\}$, with a desired distribution $P$. Let $a_1, \ldots, a_U$ be the different $U$ values of the $n$ amplitudes, and donate $n_a$ be the number of the symbol whose amplitude value is $a_u$, as described in [22], the input length $q$ of CCDM is:

$$q = \left\lfloor \log_2 \frac{n!}{U \prod_{u=1}^{n_a} n_u!} \right\rfloor,$$

where $\lfloor \cdot \rfloor$ means floor function.

Then amplitude label $\beta(\cdot)$ is utilized to map the non-equiprobable amplitudes to $b_G$ bits, which will be used to represent the relative location or amplitudes of the constellation points in the quadrant. Another $\gamma n$ equiprobable bits $b_1^n, \ldots, b_\gamma^n$ combine these non-equiprobable bits to form $h_1^n, \ldots, h_\gamma^n, \beta(A_1^j), \ldots, \beta(A_n^j)$ systematic bits of the forward error correction (FEC) encoder. If the code rate is $r$, then as defined in [22]:

$$\gamma = Z - (1-r) \log_2 M,$$

and the equivalent code rate of the system is

$$r_e = \frac{\gamma + \gamma n}{n \log_2 M}.$$

The vector $b_Q$ which represents quadrant is formed by the fraction bits $b_1^n, \ldots, b_\gamma^n$ and check bits. The check bits can be obtained by multiplying the systematic bits with the parity matrix $P$, which is a part of the generator matrix $[I|P]$ with size $(\log_2 M - Z + \gamma) n \times n \log_2 M$, and $I$ is the identity matrix.

**IV. THE AMI ANALYSIS AND THE JOINT BBPSO ALGORITHM BASED ON MAXIMIZING THE AMI**

**A. THE AMI ANALYSIS**

The AMI reflects the maximum information rate that can be transmitted without error under the given conditions. It takes multiple factors into account synthetically such as the codebook, operation SNR, channel model, etc. After setting the remaining parameters, the performance of the codebook determines the value of the AMI, so we jointly optimize the geometric shaping parameters and probability distribution of the codebook by the joint BBPSO algorithm based on maximizing the AMI.

The formulated optimization problem can be written as

$$\max \left\{ \frac{1}{K} I(X; y|H), \mu_1, \ldots, \mu_{N-1}, P_{a_1}, \ldots, P_{a_U} \right\},$$

subject to

$$s.t. \alpha_{z,j} > 1, \quad z = 1, \ldots, Z, \quad j = 1, \ldots, \sqrt{M}/2 - 1, \quad (21a)$$

$$\mu_i > 1, \quad i = 1, \ldots, N - 1, \quad (21b)$$

$$1 > p_{a_u} > 0, \quad u = 1, \ldots, U. \quad (21c)$$

where $X = [x_1, x_2, \ldots, x_J]$ is the transmitted symbols matrix constituted by all users and $H = [h_1, h_2, \ldots, h_J]$ is the corresponding channel matrix. $\alpha_{z,j}$ in (21a) is the amplitude parameter of the $z$-th $\sqrt{M}$/PAM constellation. $\mu_i$ in (21b) is the scaling parameter of the $(i+1)$-th dimension in the mother codebook. $p_{a_u}$ in (21c) is the probability of constellation points whose amplitudes are $a_u$, and eq.(21d) indicates that the sum of these probabilities is 1. $SE = c$ in eq.(21e) indicates that the optimization process is based on a predetermined SE.

According to [20], [23], the AMI of per unit of bandwidth can be written as (22), which can be found in the following page, where $K$ is the number of subcarriers, $p_m$ is the probability of the $m$-th codeword of the non-equiprobable codebook, $J$ is the number of users, $M$ is the codebook size and $\mathbb{E} [\cdot]$ denotes expectation. $P(X)$ in (22) is a modification to the AMI of non-equiprobable SCMA codebook, and

$$P(X) = \prod_{j=1}^{J} p(x_j = CB_j^m),$$

where $p(x_j = CB_j^m)$ is the probability of $x_j$ corresponding to the $m$-th codeword $CB_j^m$ of codebook $CB_j$. For equiprobable SCMA codebooks in [5] and [16], the probabilities of (23) are equal and needless when computing AMI, besides, $H(X) = \log_2(M^J)$ in this case.

The conditional probability is

$$P(y|X, H) = \frac{1}{(2\pi \sigma^2)^K} \exp\left(\sum_{k=1}^{K} -\frac{|y_k - \sum_{j=1}^{J} h^k_j x^j|^2}{2\sigma^2}\right),$$

where $y_k$ is the received signal at the $k$-th carrier, $x^j_k$ is the symbol of the $j$-th user and the $k$-th subcarrier and $h^k_j$ is the corresponding channel factor.

**B. THE JOINT BBPSO ALGORITHM BASED ON MAXIMIZING THE AMI**

For the proposed joint optimization algorithm for the PS and GS based uplink SCMA codebook, the foremost problem is how to determine the geometric shaping parameters and the probability distribution of non-uniformly distributed codewords which can achieve the highest performance gain. The optimal results depend on a number of factors such as codebook sizes, channel conditions, and so on. A scheme to achieve the geometric shaping parameters and the probability distribution by the BBPSO algorithm based on maximizing the AMI is introduced here.

BBPSO algorithm is a simple version of PSO [24] algorithm, which takes a Gaussian distribution to simplify the searching space. It can be usually described as a swarm of vectors whose trajectories oscillate around a region defined...
by each individual’s previous best optimal value and the optimal value of some other particles. Many schemes have been utilized to identify that the individual is influenced by the “some other particles” [25]. Fig.3 shows two common methods, which are gbest and lbest. In the gbest method, each particle’s search trajectory is influenced by the entire population, while the individual is only influenced by a small number of particles in the lbest method. As shown in Fig.3(b), in the typically lbest method, the particles comprise only two neighbors. It has been described that lbest populations can avoid local optima, as subpopulations explore different regions [25], and Kennedy [26] proved that populations with fewer neighbors can perform better on multimodal problems.

In order to avoid falling into local optima, the topology structure of Fig.3(b) is adopted to realize maximizing the AMI based on BBPSO algorithm. $L_4$ is designed as a set that comprises the particle itself and its two neighbors. In this case, besides the information obtained from its own search, each particle can only get their neighbor particles’ search information. In multimodal problems, when a particle $i$ finds an optimal value, the particles in $L_i$ will converge to the value quickly, but other particles can continue to explore freely. Therefore, to a certain extent, the population search is prevented from falling into the local optima. The updating equations for the $i$-th particle in the $(t+1)$-th iteration are:

$$q_{ij}(t+1) = u_{ij} + \sigma_j(t)N(0, 1)$$
$$u_{ij} = 0.5(P_{ij}(t) + T_{ij}(t))$$
$$\sigma_j(t) = |P_{ij}(t) - T_{ij}(t)|,$$

where $q_{ij}$ is the $j$-th dimension of the $i$-th particle, $i \in \{1, \ldots, W_i\}$, $j \in \{1, \ldots, D\}$, $W_i$ is the size of the swarm and $D$ is the dimensions of the search space. $N(0, 1)$ represents a standard Gaussian random variable with a mean of 0 and a variance of 1. $P_{ij}$ is the personal best for the $j$-th dimension of the $i$-th particle. $T_{ij}$ is the best position for dimension $j$ of particles in set $L_i$. The pseudo code is given by Algorithm 1.

The parameters for geometric shaping and the optimal probability distribution can be obtained by maximizing the AMI based on the joint BBPSO algorithm. According to the discussion above, the AMI depends on SNR, codebook, and so on. Firstly the relationship between the AMI and the factors of codebook is investigated in the iterative search process for a pre-set SNR. For the case of codebook size $M = 4$ under i.i.d. Rayleigh channel, there are only two different amplitude probabilities $p_{a1}$ and $p_{a2}$. Considering eq.(17), eq.(19) and $p_{a1} + p_{a2} = 1$, for a given SE, the probability distribution is determined, i.e., $p_{a1} = 0.8858$ for SNR = 1.5bps/Hz, and $p_{a1} = 0.8995$ for SE = 1.0bps/Hz. Then only the geometric shaping parameters need to be optimized. In this case, a medium SNR = 4.0dB is selected, then the AMI value is about half of the extreme value, and the optimization results in this case are proven to be suitable for most conditions. When $M = 16$, for the given SE = 3.0bps/Hz, probability distribution and geometric shaping parameters need to be jointly optimized under a pre-set SNR = 13.5dB. Then the iterative optimization algorithm for SCMA codebook is carried out to achieve the geometrical shaping parameters and the optimal probability distribution, and the convergence process are depicted in Fig.4. It can be obtained from the figure that after 10 iterations, the algorithm converges for $M = 4$ and the optimization results are $RS_1 : \alpha_{1,1} = 3.0367$, $\mu_1 = 1.0309$. When $M = 16$, the number of iterations required for convergence is 14, and the optimization results are $RS_2 : \alpha_{1,1} = 2.1838$, $\alpha_{2,1} = 1.4624$, $\mu_1 = 1.1552$, $p_{a1} = 0.3888$, $p_{a2} = 0.3423$, $p_{a3} = 0.2587$, $p_{a4} = 0.0102$.

By utilizing Monte Carlo simulation method, we compute the approximated numerical results of AMI for our proposed SCMA codebook and the codebooks in [5], [16] and LDS, and the simulation parameters are listed in the Table 1.

Fig.5 and Fig.6 describe the case with overloading factor $\eta = 150\%$, codebook size $M = 4$, in this case, for i.i.d. Rayleigh channel, when the AMI is 1.5bps/Hz (the AMI value here and below are corresponding to the simulations in section V.), as shown in Fig.5, our proposed codebook can obtain 0.6dB, 1.0dB and 1.3dB SNR gains as compared with the reference codebook [16], the original codebook [5] and LDS respectively. Fig.6 shows that the gains are 0.6dB, 0.9dB and 1.2dB respectively in tapped delay line (TDL)-C-300ns fading channel [27] when the AMI is 1.0bps/Hz.

$$\frac{1}{K} I(X; y|H) = \frac{1}{K} (H(X) - H(X|y, H)) = \frac{1}{K} (J \cdot H(P) - H(X|y, H))$$
$$= \frac{1}{K} \left( J \sum_{m=1}^{M} p_m \log_2 \frac{1}{p_m} - H_{X,H,n} \left[ \log_2 \frac{\sum_{X \in CB^j} P(X)P(y|X, H)}{P(X)P(y|X, H)} \right] \right).$$

(22)
Fig.7 and Fig.8 describe the case with overloading factor $\eta = 150\%$, codebook size $M = 16$, in the case of i.i.d. Rayleigh channel, when the AMI is 3.0bps/Hz, as shown in Fig.7, our proposed codebook can obtain 1.8dB, 2.8dB and 3.6dB SNR gains as compared with the reference codebook [16], the original codebook [5] and LDS respectively.
The gains in TDL-C-300ns fading channel are 1.4dB, 2.2dB and 2.9dB respectively, as shown in Fig.8. Specifically, as shown in the simulation results, the proposed codebook immensely narrowed the gap to the Shannon limit in low SNR region. In the high SNR region, the constellation constrained AMI of the non-equiprobable probability distribution is indeed inferior to the one of equiprobable probability distribution because the maximal entropy could be achieved if and only if the inputs are distributed uniformly [28].

V. SIMULATION RESULTS

In this section, the software simulations are carried out to assess the block error rate (BLER) performances of our proposed codebook, the reference codebook [16], the original codebook [5] and LDS, and the brief description of the codebooks in [5], [16] can be found in section II. The BLER performances are calculated for the average BLER of all users. Following are the simulation conditions. We take low density parity check (LDPC) codes as the channel code. In order to prove the robustness of our proposed scheme, we adopt different SEs in the simulations, corresponding to different code lengths \( L \), different equivalent code rates \( r_e \), different codebook sizes \( M \) and different overloading factors \( \eta \). Given \( M = 4 \), we simulate two overloading with \( \eta = 150\% \) and \( \eta = 200\% \). For i.i.d.Rayleigh channel, \( L = 800 \), \( r_e = 1/2 \) are adopted, and \( L = 900 \), \( r_e = 1/3 \) are adopted for TDL-C-300ns channel. When \( M = 16 \), \( \eta = 150\% \) is simulated, and \( L = 800 \), \( r_e = 1/2 \) are adopted for both i.i.d.Rayleigh and TDL-C-300ns channels. The Log-MPA multi-user detection algorithm is utilized, the number of inner-loop iterations is 3 and the number of iterative detection and decoding is 4, and other parameters can be found in Table 1.

When overloading factor \( \eta = 150\% \), codebook size \( M = 4 \), the SNR gains of our proposed codebook are shown in Fig.9 and Fig.10. In i.i.d.Rayleigh channel, as shown in Fig.9, when the BLER is \( 10^{-2} \), our proposed codebook can obtain 0.6dB and 0.7dB SNR gains as compared with the reference codebook [16] and the original codebook [5] respectively. The gains in TDL-C-300ns fading channel are 0.5dB and 0.65dB respectively, which can be gotten from Fig.10.

For overloading factor \( \eta = 200\% \), codebook size \( M = 4 \), the SNR gains of our proposed codebook are shown in Fig.11 and Fig.12. In i.i.d.Rayleigh channel, as shown in Fig.11, when the BLER is \( 10^{-2} \), our proposed codebook can obtain 1.4dB, 1.6dB and 1.9dB SNR gains as compared with the reference codebook [16], the original codebook [5] and LDS respectively. The gains in TDL-C-300ns fading channel are 0.9dB, 1.3dB and 1.8dB respectively, which can be gotten from Fig.12.

Fig.13 and Fig.14 describe the case with overloading factor \( \eta = 150\% \), codebook size \( M = 16 \). In the case of i.i.d.Rayleigh channel, when the BLER is \( 10^{-2} \), as shown in Fig.13, our proposed codebook can obtain 1.7dB, 2.1dB and 2.8dB SNR gains as compared with the reference...
As shown from Fig. 5 to Fig. 14, simulation results accord well with the AMI analysis, and also verify that, for different overloading, different codebook sizes, different channels and different SEs, our proposed SCMA codebooks can get significant gains, which indicates that our codebooks have strong robustness. Besides, by comparing the gains of different overloading and codebook sizes, we find that with the increase of overloading or codebook sizes, the gains are greater.

VI. CONCLUSION

In this paper, an optimization algorithm of PS and GS based codebooks for uplink SCMA scheme is proposed. This algorithm utilizes the BBPSO algorithm based on maximizing the AMI, which is simple, efficient and robust for different overloading, different channels and different codebook sizes. What’s more, with the increase of overloading or codebook sizes, the gains of our codebooks over others are greater. Based on the non-equiprobable codebook, the Log-MPA multi-user detection algorithm with non-equiprobable distribution is introduced, which makes full use of the prior information brought by the non-equiprobable signal distribution. AMI performances of our proposed codebooks, the reference codebooks are analyzed in different scenarios, which turns out that the proposed codebooks are much superior to others. Simulation results accord well with the AMI analysis, the gain is up to 2.1 dB in terms of BLER performance.

REFERENCES

[1] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," IEEE J. Sel. Areas Commun., vol. 35, no. 10, pp. 2181–2195, Oct. 2017.

[2] H. Nikopour and H. Baligh, "Sparse code multiple access," in Proc. PIMRC, London, U.K., Sep. 2013, pp. 332–336.

[3] W. Yuan, N. Wu, A. Zhang, X. Huang, Y. Li, and L. Hanzo, "Iterative receiver design for FTN signaling aided sparse code multiple access," IEEE Trans. Wireless Commun., vol. 19, no. 2, pp. 915–928, Feb. 2020.

[4] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," IEEE Trans. Signal Process., vol. 56, no. 4, pp. 1616–1626, Apr. 2008.

[5] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, "SCMA codebook design," in Proc. IEEE 80th Veh. Technol. Conf. (VTC-Fall), Vancouver, BC, Canada, Sep. 2014, pp. 1–5.

[6] N. S. Loghin, J. Zollner, B. Mouchouche, D. Ansorregui, J. Kim, and S.-I. Park, "Non-uniform constellations for ATSC 3.0," IEEE Trans. Broadcast., vol. 62, no. 1, pp. 197–203, Mar. 2016.

[7] J. Barrueco, J. Montalban, C. Regueiro, M. Velez, J. L. Ordiales, H.-M. Kim, S.-I. Park, and S. Kwon, "Constellation design of bit-interleaved coded modulation (BICM) systems in advanced broadcast standards," IEEE Trans. Broadcast., vol. 63, no. 4, pp. 603–614, Dec. 2017.

[8] Z. Qu and I. B. Djordjevic, "On the probabilistic shaping and geometric shaping in optical communication systems," IEEE Access, vol. 7, pp. 21454–21464, 2019.

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

[10] F. Steiner, G. Böcherer, and G. Liva, "Bit-metric decoding of non-binary LDPC codes with probabilistic amplitude shaping," IEEE Commun. Lett., vol. 22, no. 11, pp. 2210–2213, Nov. 2018.

[11] M. Pikus and W. Xu, "Bit-level probabilistically shaped coded modulation," IEEE Commun. Lett., vol. 21, no. 9, pp. 1929–1932, Sep. 2017.

[12] S. Zhang, K. Xiao, B. Xiao, Z. Chen, B. Xia, D. Chen, and S. Ma, "A capacity-based codebook design method for sparse code multiple access systems," in Proc. 8th Int. Conf. Wireless Commun. Signal Process. (WCSP), Yangzhou, China, Oct. 2016, pp. 1–5.
C. Jiang, Y. Wang: Uplink SCMA Codebook Design Combining PS and GS

[13] D. Cai, P. Fan, X. Lei, Y. Liu, and D. Chen, “Multi-dimensional SCMA codebook design based on constellation rotation and interleaving,” in Proc. IEEE 83rd Veh. Technol. Conf. (VTC Spring), Nanjing, China, May 2016, pp. 1–5.

[14] J. Bao, Z. Ma, Z. Ding, G. K. Karagiannidis, and Z. Zhu, “On the design of multiuser codebooks for uplink SCMA systems,” IEEE Commun. Lett., vol. 20, no. 10, pp. 1920–1923, Oct. 2016.

[15] L. Yu, X. Lei, P. Fan, and D. Chen, “An optimized design of SCMA codebook based on star-QAM signaling constellations,” in Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP), Nanjing, China, Oct. 2015, pp. 1–5.

[16] L. Yu, P. Fan, D. Cai, and Z. Ma, “Design and analysis of SCMA codebook based on star-QAM signaling constellations,” IEEE Trans. Veh. Technol., vol. 67, no. 11, pp. 10543–10553, Nov. 2018.

[17] J. Peng, W. Chen, B. Bai, X. Guo, and C. Sun, “Joint optimization of constellation with mapping matrix for SCMA codebook design,” IEEE Signal Process. Lett., vol. 24, no. 3, pp. 264–268, Mar. 2017.

[18] K. Xiao, B. Xia, L. Ma, and Y. Wei, “Design and analysis of probabilistic shaping scheme for uplink nonorthogonal multiple access systems,” IEEE Trans. Commun., vol. 67, no. 12, pp. 8818–8833, Dec. 2019.

[19] J. Kennedy, “Bare bones particle swarms,” in Proc. IEEE Swarm Intell. Symp., Apr. 2003, pp. 80–87.

[20] Q. Xie, J. Song, K. Peng, F. Yang, and Z. Wang, “Coded modulation with signal space diversity,” IEEE Trans. Wireless Commun., vol. 10, no. 2, pp. 660–669, Feb. 2011.

[21] P. Schulte and G. Bocherer, “Constant composition distribution matching,” IEEE Trans. Inf. Theory, vol. 62, no. 1, pp. 430–434, Jan. 2016.

[22] F. Buchali, F. Steiner, G. Bocherer, L. Schmalen, P. Schulte, and W. Idler, “Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration,” J. Lightw. Technol., vol. 34, no. 7, pp. 1599–1609, Apr. 1, 2016.

[23] C. Jiang and Z. Wu, “A novel uplink NOMA scheme based on low density superposition modulation,” in Proc. IEEE 86th Veh. Technol. Conf. (VTC-Fall), Sep. 2017, pp. 1–5.

[24] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in Proc. IEEE ICNN, vol. 4, Nov./Dec. 1995, pp. 1942–1948.

[25] J. Kennedy and R. Mendes, “Population structure and particle swarm performance,” in Proc. Congr. Evol. Comput. (CEC), Honolulu, HI, USA, May 2002, pp. 1671–1676.

[26] J. Kennedy, “Small worlds and mega-minds: Effects of neighborhood topology on particle swarm performance,” in Proc. Congr. Evol. Comput. (CEC), Washington, DC, USA, 1999, pp. 1931–1938.

[27] 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Study on Channel Model for Frequency Spectrum Above 6 GHz; (Release 14), Version 14.2.0, document 3GPP TR38.900, Dec. 2016.

[28] R. W. Yeung, Information Theory and Network Coding. New York, NY, USA: Springer, 2008.

CHENGXIN JIANG (Member, IEEE) received the B.S. degree in information and computing science from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2013. He is currently pursuing the Ph.D. degree in information and communication engineering with the Beijing University of Posts and Telecommunications, Beijing, China. His main research interests focus on the physical layer wireless communication and broadcasting technologies, especially the information theory, channel coding and modulation, and non-orthogonal multiple access (NOMA).

YAFENG WANG (Senior Member, IEEE) received the B.Sc. degree from the Baoji University of Arts and Science, in 1997, the M.Eng. degree from the University of Electronic Science and Technology of China, in 2000, and the Ph.D. degree from the Beijing University of Posts and Telecommunications, in 2003. In 2008, he was a Visiting Scholar with the Faculty of Engineering and Surveying, University of Southern Queensland, Australia. He is currently a Professor of electronic engineering with the School of Information and Telecommunications, Beijing University of Posts and Telecommunications. He leads the Broadband Mobile Communication Engineering Laboratory, which is one of Zhong Guan Cun Science Park Open Laboratories. He has authored or coauthored over 100 peer-reviewed journal articles and conference papers. His research mainly focuses on wireless communications and information theory.