

Design of SCMA Codebooks Using Differential Evolution

Kuntal Deka¹, Minerva Priyadarsini¹, Sanjeev Sharma², and Baltasar Beferull-Lozano (Senior Member, IEEE)³

¹Indian Institute of Technology Goa, India, ²Indian Institute of Technology (BHU) Varanasi, India

³Department of Information and Communication Technology, University of Agder, Grimstad 4879, Norway

Abstract—Non-orthogonal multiple access (NOMA) is a promising technology which meets the demands of massive connectivity in future wireless networks. Sparse code multiple access (SCMA) is a popular code-domain NOMA technique. The effectiveness of SCMA comes from: (1) the multi-dimensional sparse codebooks offering high shaping gain and (2) sophisticated multi-user detection based on message passing algorithm (MPA). The codebooks of the users play the main role in determining the performance of SCMA system. This paper presents a framework to design the codebooks by taking into account the entire system including the SCMA encoder and the MPA-based detector. The symbol-error rate (SER) is considered as the design criterion which needs to be minimized. Differential evolution (DE) is used to carry out the minimization of the SER over the codebooks. The simulation results are presented for various channel models.

Index Terms—SCMA, codebook design, differential evolution, message passing algorithm.

I. INTRODUCTION

The future wireless communication system will comprise of enormous number of interconnected devices. The massive connectivity required in such situations cannot be fulfilled by the multiple access schemes deployed so far. Earlier, 3G system used code division multiple access (CDMA), 4G used orthogonal frequency division multiple access (OFDMA) [1]. All these technologies are based on orthogonal multiple access (OMA) principle. In OMA, the number of supported users is limited by the number of available orthogonal resources. The non-orthogonal multiple access (NOMA) technology provides a gateway for massive connectivity. It supports overloaded systems where the number of users is higher than the number of orthogonal resources. NOMA techniques can be broadly classified into two groups: *power domain* and *code domain*. In power domain NOMA, different levels of powers are allocated to different users. In code domain NOMA, different code-words or signatures are used for different users. Low-density spreading (LDS) is a code-domain NOMA technique where, a user's symbol is multiplied with a distinct sparse spreading signature. This spread-ed sequence is mapped to modulation constellation points for transmission [2]. Nikoupor *et al.* came up with the technique of sparse code multiple access (SCMA) in order to improve upon LDS [3]. In SCMA, the operations of the spreading and the modulation mapping are merged. Based on a dedicated codebook, a symbol is directly mapped to a sparse multi-dimensional codeword. Thus, SCMA provides a better opportunity than LDS to attain high shaping gain. Owing to the inherent sparsity in the code-domain NOMA, the multi-user detection is usually done by the message passing algorithm (MPA) [4].

Related work: The performance of an SCMA system is mainly determined by the codebooks of the users. The optimization of the codebooks for SCMA system is a convoluted

task as multiple users are interfering with multi-dimensional complex vectors. In [3, 5], first, the optimum codebook for one user is designed. The remaining codebooks were obtained by carrying out user-specific operations on the optimized codebook. In [6] (referred to as “Zhang [6]”), the sum-rate was considered as one of the design criteria. First, a series of one-dimensional complex codewords were designed. Then the angles of these codewords were changed with the objective of improving the sum-rate. The authors in [7] considered the minimum Euclidean distance and the energy diversity of the constellation points to obtain the optimum codebooks. Yu *et al.* (referred to as “Yu [8]”) designed SCMA codebooks based on the star quadrature amplitude modulation (QAM) constellations [8]. In [9], the authors designed codebooks for bit-interleaved convolutionally-coded SCMA system. This design was aided by the analysis based on EXtrinsic Information Transfer (EXIT) chart. Sharma *et al.* [10] (referred to as “Sharma [10]”) designed the codebooks by maximizing the mutual information and shaping gain.

Contributions: The existing codebook-design methods mainly focus on various geometric properties of the multi-dimensional constellations with little emphasis on detection. It is difficult to track the MPA-based detection process mathematically. The factor graph is finite, not tree-like and it contains short cycles. Due to these reasons, the techniques like density evolution and EXIT chart cannot accurately characterize the MPA-based multi-user detection process. We propose to consider the symbol error rate (SER) as the cost function which is to be minimized. The SER is one such quantity which takes into account every part of the multi-user system be it Euclidean distance profile, product distance profile, MPA etc. The reliability of the system is precisely reflected by the SER. We adopt differential evolution (DE) for the minimization of the SER as it is not a simple function of the codebooks. DE is a flexible and effective evolutionary algorithm which is used to solve complex optimization problems with real-valued parameters [11, 12]. First, the structure of the codebooks is represented with the help of a finite number of constellation points. Then a DE-based optimization process is invoked to find the optimum constellation points. The codebooks for the additive white Gaussian noise (AWGN) and Rayleigh fading channels are designed. The SER performance of the proposed codebooks are compared with those of the existing ones. Moreover, various key parameters of the codebooks are computed and analyzed.

Outline: Section II describes the preliminaries such as SCMA system model, important parameters of codebooks and DE. The proposed method of codebook design is presented in Section III. The simulation results are presented and analyzed

in Section IV. Section V concludes the paper.

II. PRELIMINARIES

A. System Model

A $J \times K$ SCMA system refers to an overloaded multi-user scenario with J users and K resource elements. The overloading factor is given by $\lambda = \frac{J}{K}$. A dedicated codebook \mathcal{C}_j containing M K -dimensional codewords: $\mathcal{C}_j = \{\mathbf{x}_{j1}, \mathbf{x}_{j2}, \dots, \mathbf{x}_{jM}\}$ is assigned to every j^{th} user. Each codeword \mathbf{x}_{jm} is sparse with N non-zero complex components. Based on the assigned codebook \mathcal{C}_j , $\log_2(M)$ data bits are directly mapped to the K -dimensional codeword $\mathbf{x}_j = [x_{j1}, \dots, x_{jK}]^T$. The received signal $\mathbf{y} = [y_1, \dots, y_K]^T$ is:

$$\mathbf{y} = \sum_{j=1}^J \text{diag}(\mathbf{h}_j) \mathbf{x}_j + \mathbf{n} \quad (1)$$

where, $\mathbf{h}_j = [h_1, \dots, h_K]^T$ is the channel gain vector for the j^{th} user and \mathbf{n} is a complex $K \times 1$ AWGN vector.

The locations of the non-zero elements of the codewords for the users can be represented with the help of a factor matrix as shown in (2). The 1s present in the j^{th} column specify the locations of the non-zero components of the codewords for the j^{th} user. The matrix F can be alternatively represented by a factor graph as shown in Fig. 1. The degree of a resource node is denoted by d_f . This specifies that d_f users interfere with each others over one resource element. As the factor graph is sparse, MPA is used for multi-user detection.

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \quad (2)$$

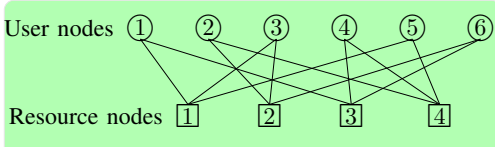


Fig. 1: Factor graph of six users ($J = 6$) and four resource nodes ($K = 4$) where three users overlap over one resource.

The performance of an SCMA system is highly sensitive to the codebooks. In literature [13, 14], various key performance indicators (KPIs) have been considered for the design of multi-dimensional constellations. A few of them are highlighted:

- **Minimum Euclidean Distance** ($d_{E,\min}$): It is defined as the minimum distance between any pair ($\mathbf{x}_m, \mathbf{x}_n$) of codewords in the entire SCMA system:

$$d_{E,\min} = \min_{m,n} \|\mathbf{x}_m - \mathbf{x}_n\|.$$

- **Euclidean kissing number** (τ_E): It is defined as the number of distinct codeword pairs with Euclidean distance equal to $d_{E,\min}$.

- **Minimum product distance** ($d_{P,\min}$): The product distance $d_P^{m,n}$ between two codewords \mathbf{x}_m and \mathbf{x}_n is defined as

$$d_P^{m,n} = \prod_{j \in \mathcal{J}_{m,n}} |x_{mj} - x_{nj}|$$

where, $\mathcal{J}_{m,n}$ is the set of dimensions for which $x_{mj} \neq x_{nj}$. The minimum product distance $d_{P,\min}$ is given by:

$$d_{P,\min} = \min_{m,n} d_P^{m,n}.$$

- **Product Kissing Number** (τ_P): It is defined as the number of distinct codeword pairs with product distance equal to $d_{P,\min}$.

The objective is to maximize $d_{E,\min}$, $d_{P,\min}$ and minimize τ_E , τ_P . In addition to these parameters, mutual information between the received signal and the sum of the interfering codewords is also a vital parameter. Suppose Y is the signal received over one resource element. Let S represent the sum of the codewords of the d_f users interfering over that resource. We have a total of M^{d_f} distinct sum values for S as given by $\{s_1, s_2, \dots, s_{M^{d_f}}\}$. A high value of $I(Y; S)$ ensures successful recovery of the individual user's data from the noisy sum value. Usually a lower bound I_L on $I(Y; S)$ is considered¹. I_L for AWGN channel with noise variance N_0 is given by [6]

$$I_L = \log_2 M^{d_f} - \log_2 \left[1 + \frac{1}{M^{d_f}} \sum_{j=1}^{M^{d_f}} \sum_{i \neq j}^{M^{d_f}} \exp \left(-\frac{1}{4N_0} |s_j - s_i|^2 \right) \right]. \quad (3)$$

B. Differential Evolution

In DE, an initial set of S_P random candidate solutions or vectors $\{\mathbf{p}_i : i = 1, 2, \dots, S_P\}$ is generated. The length of each vector is the same and it is denoted by D . New generations are created iteratively using mutation, cross over and selection process. During mutation, each vector \mathbf{p}_i^G , where G denotes the generation index, is selected as a primary parent. For each parent, a mutant vector is generated as follows:

$$\mathbf{u}_i^G = \mathbf{p}_{r_1} + \alpha(\mathbf{p}_{r_2} - \mathbf{p}_{r_3}) \quad (4)$$

where, r_1, r_2, r_3 are randomly selected distinct numbers different from i and α is the scaling factor. \mathbf{u}_i^G is the secondary parent generated from mutation. Cross-over is applied to primary and secondary parent to obtain the offspring vector \mathbf{v}_i^G . Each component v_{ik}^G of $\mathbf{v}_i^G = [v_{i1}^G, \dots, v_{ik}^G, \dots, v_{iD}^G]$ is inherited from either the primary or the secondary parent as per the following rule:

$$v_{ik}^G = \begin{cases} u_{ik}^G, & \text{if } h_k \leq C_r \\ p_{ik}^G, & \text{otherwise} \end{cases}$$

where, h_k is a random number uniformly distributed over $[0, 1]$, C_r is the cross-over rate. The offspring is made to inherit at least one component from the secondary parent to ensure that offspring is different from primary parent. The offspring vector \mathbf{v}_i^G has to compete with the parent vector \mathbf{p}_i^G to get a place in the next $G + 1$ generation. If \mathbf{v}_i^G gives a lower cost function than \mathbf{p}_i^G then the former replaces the later, else the primary parent vector exists in the next generation too, i.e.

$$\mathbf{p}_i^{(G+1)} = \begin{cases} \mathbf{v}_i^G, & \text{if } f(\mathbf{v}_i^G) \leq f(\mathbf{p}_i^G) \\ \mathbf{p}_i^G, & \text{otherwise} \end{cases}$$

¹ $I(Y; S)$ or I_L is related to the sum-rate [6].

where $f(\cdot)$ is the objective function that needs to be minimized. The new generation performs at least as good as the best candidate of the previous generation. These steps are continued until some stopping criteria are satisfied. The best vector from the current population is considered as the optimum vector.

III. PROPOSED CODEBOOK DESIGN BASED ON DIFFERENTIAL EVOLUTION

Let $\mathbf{C} = \{\mathcal{C}_1, \dots, \mathcal{C}_J\}$ denote the set or the collection of the codebooks of all the users. The task of SCMA codebook design is formulated as an optimization problem as follows:

$$\begin{aligned} \mathbf{C}_{\text{opt}} = \arg \min_{\mathbf{C}} f\left(\mathbf{C}; \frac{E_b}{N_0}, \mathbf{F}\right) \\ \text{s.t. } \|\mathbf{c}\|_2 = 1 \quad \forall \mathbf{c} \in \mathbf{C} \end{aligned} \quad (5)$$

where, $f(\cdot)$ is the SER at SNR $\frac{E_b}{N_0}$ and \mathbf{F} is the factor graph matrix. Every K -dimensional codeword \mathbf{c} present in the SCMA codebook system \mathbf{C} is constrained to have unit Euclidean norm.

Note that an already-designed factor graph matrix \mathbf{F} is fed to the optimization process in (5). Suppose the factor graph is regular with the resource node degree of d_f . Then, d_f users overlap on a particular resource node. One needs to carefully assign constellation points to these users. The task is to assign d_f distinct one-dimensional constellation² codebooks to the d_f overlapping users. These codebooks may be represented by $C_1^o, C_2^o, \dots, C_{d_f}^o$, where the superscript 'o' signifies that these are one-dimensional. The size of the constellation is M . For example, consider the case of 6×4 SCMA system with the factor graph shown in Fig. 1 and the matrix in (2). Here, $d_f = 3$ and $M = 4$. The three one-dimensional codebooks can be represented as follows:

$$\begin{aligned} C_1^o &= [a_1 \ a_2 \ -a_2 \ -a_1] \\ C_2^o &= [a_3 \ a_4 \ -a_4 \ -a_3] \\ C_3^o &= [a_5 \ a_6 \ -a_6 \ -a_5] \end{aligned} \quad (6)$$

where, $a_i \in \mathbb{C}$, $i = 1, \dots, 6$.

In (6), 12 distinct complex numbers are created from 6 complex numbers and their negative counterparts. It is preferable to design the SCMA codebooks with the minimum possible number of constellation points. This reduces the hardware requirement in implementation.

The one-dimensional codebooks can be assigned to all the resource nodes through a structure matrix satisfying Latin property [15]. This way of generating the entire set of the codebooks was considered in [6, 16]. The structure matrix for the 6×4 SCMA system is shown below:

$$\mathbf{F}_L = \begin{bmatrix} C_1^o & 0 & C_2^o & 0 & C_3^o & 0 \\ 0 & C_2^o & C_3^o & 0 & 0 & C_1^o \\ C_2^o & 0 & 0 & C_1^o & 0 & C_3^o \\ 0 & C_1^o & 0 & C_3^o & C_2^o & 0 \end{bmatrix} \quad (7)$$

The Latin property in (7) ensures that distinct one-dimensional codebooks are assigned to the users overlapping

²One-dimensional codebook is assigned as only one resource node is considered at a time.

over any particular resource. Using \mathbf{F}_L , the set of all codebooks can be generated as follows:

$$c_1 = \begin{bmatrix} C_1^o \\ 0 \\ C_2^o \\ 0 \end{bmatrix} \quad c_2 = \begin{bmatrix} 0 \\ C_2^o \\ 0 \\ C_1^o \end{bmatrix} \quad c_3 = \begin{bmatrix} C_2^o \\ C_3^o \\ 0 \\ 0 \end{bmatrix} \quad c_4 = \begin{bmatrix} 0 \\ 0 \\ C_1^o \\ C_3^o \end{bmatrix} \quad c_5 = \begin{bmatrix} C_3^o \\ 0 \\ 0 \\ C_2^o \end{bmatrix} \quad c_6 = \begin{bmatrix} 0 \\ C_1^o \\ C_3^o \\ 0 \end{bmatrix} \quad (8)$$

where, $\mathbf{0}$ is the all-zero row vector of length $M = 4$.

In order to obtain more diversity and higher shaping gain, the authors in [6, 16] have proposed to carry out the dimensional permutation switching algorithm (DPSA). As per DPSA, the permuted version of C_2^o denoted by $C_{2,p}^o$ is used in the codebooks \mathcal{C}_3 and \mathcal{C}_5 in place of C_2^o . We consider $C_{2,p}^o = [-a_4, a_3, -a_3, a_4]$. With this dimensional permutation, using (6) in (8), the codebooks can be written as shown in TABLE I. Observe that the codebooks for the 6×4

TABLE I: Structure \mathbb{S} of the codebooks $\mathbf{C} = \{\mathcal{C}_1, \dots, \mathcal{C}_6\}$

$$\begin{aligned} \mathcal{C}_1 &= \left\{ \begin{bmatrix} a_1 \\ 0 \\ a_3 \\ 0 \end{bmatrix}, \begin{bmatrix} a_2 \\ 0 \\ a_4 \\ 0 \end{bmatrix}, \begin{bmatrix} -a_2 \\ 0 \\ -a_4 \\ 0 \end{bmatrix}, \begin{bmatrix} -a_1 \\ 0 \\ -a_3 \\ 0 \end{bmatrix} \right\} \quad \mathcal{C}_2 = \left\{ \begin{bmatrix} 0 \\ a_3 \\ 0 \\ a_1 \end{bmatrix}, \begin{bmatrix} 0 \\ a_4 \\ 0 \\ a_2 \end{bmatrix}, \begin{bmatrix} 0 \\ -a_4 \\ 0 \\ -a_2 \end{bmatrix}, \begin{bmatrix} 0 \\ -a_3 \\ 0 \\ -a_1 \end{bmatrix} \right\} \\ \mathcal{C}_3 &= \left\{ \begin{bmatrix} -a_4 \\ a_5 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} a_3 \\ a_6 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -a_3 \\ -a_6 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} a_4 \\ -a_5 \\ 0 \\ 0 \end{bmatrix} \right\} \quad \mathcal{C}_4 = \left\{ \begin{bmatrix} 0 \\ 0 \\ a_1 \\ a_5 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ a_2 \\ a_6 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -a_2 \\ -a_6 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -a_1 \\ -a_5 \end{bmatrix} \right\} \\ \mathcal{C}_5 &= \left\{ \begin{bmatrix} a_5 \\ 0 \\ 0 \\ -a_4 \end{bmatrix}, \begin{bmatrix} a_6 \\ 0 \\ 0 \\ a_3 \end{bmatrix}, \begin{bmatrix} -a_6 \\ 0 \\ 0 \\ -a_3 \end{bmatrix}, \begin{bmatrix} -a_5 \\ 0 \\ 0 \\ a_4 \end{bmatrix} \right\} \quad \mathcal{C}_6 = \left\{ \begin{bmatrix} 0 \\ a_1 \\ a_5 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ a_2 \\ a_6 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -a_2 \\ -a_6 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -a_1 \\ -a_5 \\ 0 \end{bmatrix} \right\} \end{aligned}$$

SCMA system can be represented with the help of 6 complex numbers a_i , $i = 1, \dots, 6$. These complex numbers can be compactly represented by $\mathbf{a} = [a_1, \dots, a_6]$. Any particular complex vector \mathbf{a}_i , $\mathbf{a}_i \in \mathbb{C}^6$ refers to a particular set of codebooks $\mathbf{C}_i = \{\mathcal{C}_1^i, \dots, \mathcal{C}_6^i\}$. With these notations, the SCMA codebook design problem in (5) can now be written as

$$\mathbf{a}_{\text{opt}} = \arg \min_{\mathbf{a}} f\left(\mathbf{a}; \frac{E_b}{N_0}\right) \quad (9)$$

where, $f\left(\mathbf{a}; \frac{E_b}{N_0}\right)$ is the SER of the SCMA system defined by \mathbf{a} at SNR $\frac{E_b}{N_0}$. In the rest of the paper, for notational brevity, $f\left(\mathbf{a}; \frac{E_b}{N_0}\right)$ will be represented as $f(\mathbf{a})$ with the understanding that the SNR is fixed at a particular value during the optimization process. In (9), \mathbf{a}_{opt} is that vector \mathbf{a} which yields the optimum codebook \mathbf{C}_{opt} as per the structure given in TABLE I. Moreover, we have the constraint that the Euclidean norm of every codeword is 1 although it is not explicitly mentioned in (9).

We solve (9) with the help DE. The job is to find the optimum vector \mathbf{a} which contains 6 complex numbers³. However, DE is a real-valued optimization technique. Therefore, in this case, the number of real-valued variables to be optimized is $D = 12$. The detailed steps for the DE-based codebook design

³It is noteworthy that the proposed method of codebook design is universal in the sense that it can be applied to an SCMA system with any values of J and K . The structure as shown in TABLE I must be fed to the DE-based optimizer.

Algorithm 1: Codebook design using differential evolution

input : Codebook structure \mathbb{S} (TABLE I), $\frac{E_b}{N_0}$, DE parameters (α , C_r , S_P and I_{\max})

output: Codebooks $\mathbf{C}_{\text{opt}} = \{\mathbf{C}_1, \dots, \mathbf{C}_6\}$.

Initialize the population matrix \mathbf{P} to a matrix of size $S_P \times D$ having random numbers uniformly distributed over $[-1, 1]$ where $D = 12$, Normalize these entries so that the every codeword of the codebook has unit norm;

while termination criteria not fulfilled **do**

for $i \leftarrow 1$ **to** S_P **do**

 Select three distinct vectors (rows) $\mathbf{p}_{r_0}, \mathbf{p}_{r_1}$ and \mathbf{p}_{r_2} uniformly at random from \mathbf{P} such that they are also different from \mathbf{p}_i ;

 Generate an integer j_{rand} uniformly at random from $\{1, 2, \dots, D\}$;

 /* Generation of trial vector \mathbf{u} */

for $j \leftarrow 1$ **to** D **do**

if $\text{rand}[0,1] \leq C_r$ or $j = j_{\text{rand}}$ **then**

$u_{j,i} = p_{j,r_0} + \alpha \times (p_{j,r_1} - p_{j,r_2})$

 // Crossover and Mutation

else

$u_{j,i} = p_{j,i}$

 Normalize the entries of \mathbf{u} so that every codeword has unit norm.

 /* Evaluation and Selection */

 Suppose \mathbf{C}^u and \mathbf{C}^{p_i} are the two sets of the codebooks for all users as per \mathbf{u} and \mathbf{p}_i respectively.

 Run Monte Carlo simulation for the SCMA system with the codebooks \mathbf{C}^u and \mathbf{C}^{p_i} at SNR $\frac{E_b}{N_0}$. Suppose, $f(\mathbf{u})$ and $f(\mathbf{p}_i)$ are the respective SER values ;

if $f(\mathbf{u}) < f(\mathbf{p}_i)$ **then**

$\mathbf{p}_i = \mathbf{u}$ // Replace the i th row of \mathbf{P} by \mathbf{u} ;

From the updated population matrix \mathbf{P} , find the vector (row) \mathbf{p}_{opt} which yields the minimum value of SER;

If $f(\mathbf{p}_{\text{opt}})$ is not changing significantly from the previous iteration or the maximum number of iterations I_{\max} are exhausted, then break from loop;

Using \mathbf{p}_{min} , form the optimum constellation vector

$\mathbf{a}_{\text{opt}} = [a_1, \dots, a_6]$. Then inserting \mathbf{a} in \mathbb{S} , the optimized set of the codebooks of all users $\mathbf{C}_{\text{opt}} = \{\mathbf{C}_1, \dots, \mathbf{C}_J\}$ are obtained. ;

are presented in Algorithm 1. The inputs to the algorithm are the structure \mathbb{S} for the codebooks as shown in TABLE I, SNR $\frac{E_b}{N_0}$ and the DE parameters: α , C_r , S_P and I_{\max} . The candidate vectors are stored in a population matrix \mathbf{P} in row-wise manner. The total number of candidate vectors is S_P . The size of \mathbf{P} is $S_P \times D$. The elements of \mathbf{P} are initialized to random numbers uniformly distributed over $[-1, 1]$. Every row \mathbf{p}_s , $s = 1, \dots, S_P$, of \mathbf{P} corresponds to an SCMA codebook system specified by the 6 complex numbers $\mathbf{a}_s = [a_{s,1}, \dots, a_{s,6}]$. Suppose $a_{s,t} = a_{s,t}^r + a_{s,t}^c i$, $t = 1, \dots, 6$, where $a_{s,t}^r$ and $a_{s,t}^c$ are real uniform numbers in $[-1, 1]$. The s^{th} row \mathbf{p}_s of \mathbf{P} is given by

$$\mathbf{p}_s = [a_{s,1}^r, a_{s,1}^c, a_{s,2}^r, a_{s,2}^c, \dots, a_{s,6}^r, a_{s,6}^c].$$

The elements of \mathbf{P} are normalized so that every codeword in a codebook system has unit norm. Against each row \mathbf{p}_s , $s = 1, \dots, S_P$, a trial vector \mathbf{u} is generated with the given values of the DE parameters: crossover rate (C_r) and scaling factor (α). Algorithm 1 shows the detailed steps of the

generation of the trial vector \mathbf{u} . The decision regarding the replacement of the current population vector \mathbf{p}_s by the trial vector \mathbf{u} is made by computing the SER values for the two SCMA systems through Monte Carlo simulations. If the SER value $f(\mathbf{u})$ of the SCMA system defined by \mathbf{u} is less than the SER $f(\mathbf{p}_i)$ of the system defined by \mathbf{p}_i , then the current population vector \mathbf{p}_i is set to \mathbf{u} . In this way, every candidate vector in \mathbf{P} is examined and updated with the trial vectors if necessary. From the updated population \mathbf{P} , the best vector \mathbf{p}_{opt} with the minimum SER value is identified. The above-mentioned process is repeated unless $f(\mathbf{p}_{\text{opt}})$ is not changing significantly during two consecutive iterations or the maximum number of iterations are exhausted. The real numbers in \mathbf{p}_{opt} converted to the corresponding complex numbers to obtain \mathbf{a}_{opt} . The complex numbers in $\mathbf{a}_{\text{opt}} = [a_1, \dots, a_6]$ are plugged into \mathbb{S} in TABLE I to obtain the optimum codebooks.

Example 1. Consider the case of $J = 6$, $K = 4$ SCMA system with constellation size $M = 4$. The structure given in TABLE I is used. The number of variables is $D = 12$. We can consider $S_P = 20$, $C_r = 0.95$ and $\alpha = 0.6$. The population matrix \mathbf{P} may be initialized to the following 20×12 matrix:

$$\mathbf{P} = \begin{bmatrix} 0.46 & -0.53 & -0.90 & -0.22 & -0.68 & 0.16 & -0.34 & -0.08 & 0.22 & 0.91 & 0.49 & 0.48 \\ -0.02 & 0.90 & 0.44 & 0.08 & -0.44 & 0.01 & 0.32 & -0.84 & 0.10 & 0.24 & 0.82 & 0.12 \\ -0.65 & -0.62 & 0.80 & 0.11 & -0.41 & -0.15 & -0.55 & 0.22 & 0.56 & 0.22 & 0.41 & 0.15 \end{bmatrix}.$$

Due to space constraint, only 3 rows out of 20 are shown. It may be verified that the norm of every codeword in \mathbf{P} is 1. For every row of \mathbf{P} , a trial vector is generated by carrying out the mutation and the crossover operations as specified in Algorithm 1. If the trial vector yields less SER than the current row vector, then the current row is overwritten by the trial vector. In this way every row of \mathbf{P} is examined and updated if needed. Suppose, finally, the population matrix \mathbf{P} is as given below where the first row is the best row \mathbf{p}_{min} .

$$\mathbf{P} = \begin{bmatrix} -0.33 & 0.63 & -0.83 & 0.43 & 0.71 & 0 & -0.36 & 0 & -0.42 & -0.84 & 0.59 & 0.35 \end{bmatrix}.$$

Then the solution of (9) is given by the following:

$$\mathbf{a}_{\text{opt}} = [-0.33 + 0.63i, -0.83 + 0.43i, 0.71, -0.36, -0.42 - 0.84i, 0.59 + 0.35i].$$

Using the above \mathbf{a}_{opt} in TABLE I, we can generate the codebooks for the SCMA system. TABLE II shows these codebooks in Section IV.

Remark 1. Although the proposed scheme appears to be SNR-dependent, the codebooks optimized at a high SNR are found to work well at different SNR values. In practice, we fix an SNR where the SER is around 10^{-3} . The results for different SNR values are not shown here due to space constraint.

Complexity Analysis: Observe from Algorithm 1 that during each cycle of DE, the objective functions (SERs) for the current population and the trial vector are computed. The SER computation is a time-consuming process due to fact that the complexity of the MPA is $O(M^{d_f})$. Thus, the complexity of proposed algorithm is higher than some of the existing codebook designing methods. However, note that the codebook design is a one-time and offline task. It does not add to the real-time complexity of the system. Therefore, the proposed DE-based codebook design method is feasible in practical scenario.

IV. SIMULATION RESULTS

The simulations are carried out to evaluate the performance of the proposed codebooks. For comparison, the following codebooks are considered: • “Zhang [6]” • “Sharma [10]” • “Yu [8]” • “Ma [17]” (Chapter 12 of [17]). As per the thumb rules provided in [18, 19], the DE parameters are set to: $S_P = 20$, $D = 12$, $C_r = 0.95$, $\alpha = 0.6$ and $I_{\max} = 80$. The simulations are done for uncoded transmission over AWGN and flat Rayleigh fading channels.

A. AWGN channel

First the results for the AWGN channel are presented. An SCMA system with $J = 6$, $K = 4$ and $M = 4$ is considered. Algorithm 1 is run at $\frac{E_b}{N_0} = 10$ dB to find the optimum vector \mathbf{a}_{opt} . The complex numbers of \mathbf{a}_{opt} and their negative versions are shown in Fig. 2.

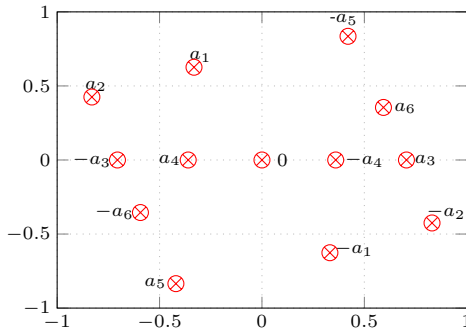


Fig. 2: \mathbf{a}_{opt} for AWGN channel

With this value of $\mathbf{a}_{\text{opt}} = [a_1, a_2, a_3, a_4, a_5, a_6]$, the optimum codebooks are generated using the structure \mathbb{S} shown in TABLE I. These codebooks are shown in TABLE II. Observe

TABLE II: Codebooks optimized for AWGN channel

$\mathbf{c}_1 = \begin{bmatrix} -0.3318 + 0.6262i \\ 0 \\ 0.7055 \\ 0 \end{bmatrix}$	$\begin{bmatrix} -0.8304 + 0.4252i \\ 0 \\ -0.3601 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0.8304 - 0.4252i \\ 0 \\ 0.3601 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0.3318 - 0.6262i \\ 0 \\ -0.7055 \\ 0 \end{bmatrix}$
$\mathbf{c}_2 = \begin{bmatrix} 0 \\ 0.7055 \\ 0 \\ -0.3318 + 0.6262i \end{bmatrix}$	$\begin{bmatrix} 0 \\ -0.3601 \\ 0 \\ -0.8304 + 0.4252i \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0.3601 \\ 0 \\ 0.8304 - 0.4252i \end{bmatrix}$	$\begin{bmatrix} 0 \\ -0.7055 \\ 0 \\ 0.3318 - 0.6262i \end{bmatrix}$
$\mathbf{c}_3 = \begin{bmatrix} 0.3601 \\ -0.4202 - 0.8350i \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0.7055 \\ 0.5933 + 0.3548i \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} -0.7055 \\ -0.5933 - 0.3548i \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} -0.3601 \\ 0.4202 + 0.8350i \\ 0 \\ 0 \end{bmatrix}$
$\mathbf{c}_4 = \begin{bmatrix} 0 \\ 0 \\ -0.3318 + 0.6262i \\ -0.4202 - 0.8350i \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ -0.8304 + 0.4252i \\ 0.5933 + 0.3548i \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0.8304 - 0.4252i \\ -0.5933 - 0.3548i \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0.3318 - 0.6262i \\ 0.4202 + 0.8350i \end{bmatrix}$
$\mathbf{c}_5 = \begin{bmatrix} -0.4202 - 0.8350i \\ 0.5933 + 0.3548i \\ 0 \\ 0.3601 \end{bmatrix}$	$\begin{bmatrix} 0.7055 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} -0.7055 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} -0.3601 \\ 0 \\ 0 \\ 0 \end{bmatrix}$
$\mathbf{c}_6 = \begin{bmatrix} -0.3318 + 0.6262i \\ -0.4202 - 0.8350i \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} -0.8304 + 0.4252i \\ 0.5933 + 0.3548i \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0.8304 - 0.4252i \\ -0.5933 - 0.3548i \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0.3318 - 0.6262i \\ 0.4202 + 0.8350i \\ 0 \\ 0 \end{bmatrix}$

that Euclidean norm of every codeword is 1.

The SER performance of the proposed codebooks are evaluated through Monte Carlo simulations and these are shown in Fig. 3. The SER plots for the other codebooks are also presented. Observe that the proposed codebooks outperform

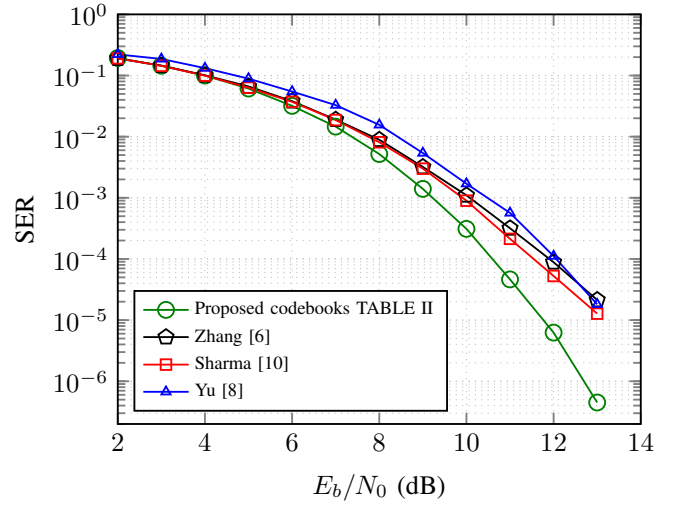


Fig. 3: SER performance of the SCMA system using various codebooks for $J = 6$ and $K = 4$ in AWGN channel.

the others by a significant margin. Specifically, there is coding gain of about 1.4 dB at $\text{SER}=10^{-5}$ over the next best codebooks (“Sharma [10]”). The performance of “Ma [17]” over AWGN channel is not satisfactory. Therefore, its SER plot is not shown in Fig. 3. However, “Ma [17]” yields good results over Rayleigh fading channel which is presented later in this section.

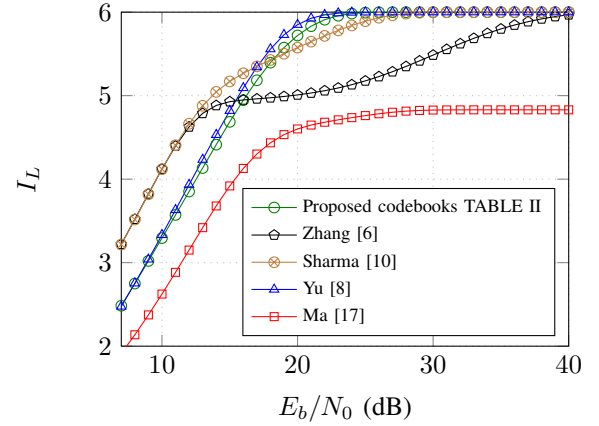


Fig. 4: Lower bound on the mutual information for various codebooks.

The SER performance of the various codebooks can be studied and justified by carrying out the mutual information analysis as mentioned in Section II-A. We plot the lower bound I_L on mutual information for various codebooks in Fig. 4. Observe that in the SNR region below 16 dB, “Sharma [10]” and “Zhang [6]” provide higher values of I_L than the proposed codebooks. However, the I_L for “Yu [8]” and the proposed codebooks reach the maximum value of 6 quicker than the other codebooks. Observe that I_L for “Yu [8]” is slightly higher than that for the proposed codebooks in the range of 15-20 dB. They reach the maximum value almost at the same SNR of 24 dB. However, as described later in this section, the Euclidean distance and the product distance

profiles of “Yu [8]” are poorer than those of the proposed one. This observation reinforces the superior performance of the proposed DE-based codebooks. Also note that the I_L for “Ma [17]” cannot climb up to the maximum value. This justifies the poor performance of “Ma [17]” over the AWGN channel.

B. Fading channel

In this case, each user observes independent Rayleigh fading channel coefficients over the resource elements. The same SCMA framework with $J = 6$, $K = 4$, $M = 4$ and the structure \mathbb{S} given in TABLE I is considered. Algorithm 1 is executed at $\frac{E_b}{N_0} = 17$ dB. It yields the optimum constellation vector \mathbf{a}_{opt} which are plotted in Fig. 5. The complex numbers

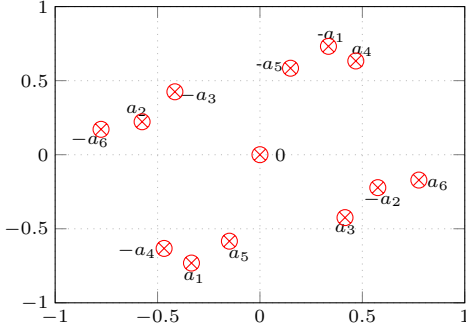


Fig. 5: \mathbf{a}_{opt} for fading channel

in $\mathbf{a}_{\text{opt}} = [a_1, \dots, a_6]$ are put in \mathbb{S} . The resulting codebooks are shown in TABLE III. Observe that the optimized codebooks for the fading channel are different from those for the AWGN channel. This signifies that the codebook design problem depends on the underlying channel model.

TABLE III: Codebooks optimized for fading channel

$$\begin{aligned}
 \mathbf{c}_1 &= \begin{Bmatrix} \begin{bmatrix} -0.3344 - 0.7316i \\ 0 \\ 0.4153 - 0.4248i \\ 0 \end{bmatrix} & \begin{bmatrix} -0.5754 + 0.2224i \\ 0 \\ 0.4680 + 0.6328i \\ 0 \end{bmatrix} & \begin{bmatrix} 0.5754 - 0.2224i \\ 0 \\ -0.4680 - 0.6328i \\ 0 \end{bmatrix} & \begin{bmatrix} 0.3344 + 0.7316i \\ 0 \\ -0.4153 + 0.4248i \\ 0 \end{bmatrix} \\
 \mathbf{c}_2 &= \begin{Bmatrix} \begin{bmatrix} -0.3344 - 0.7316i \\ 0 \\ 0.4153 - 0.4248i \\ 0 \end{bmatrix} & \begin{bmatrix} -0.5754 + 0.2224i \\ 0 \\ 0.4680 + 0.6328i \\ 0 \end{bmatrix} & \begin{bmatrix} 0.5754 - 0.2224i \\ 0 \\ -0.4680 - 0.6328i \\ 0 \end{bmatrix} & \begin{bmatrix} 0.3344 + 0.7316i \\ 0 \\ -0.4153 + 0.4248i \\ 0 \end{bmatrix} \\
 \mathbf{c}_3 &= \begin{Bmatrix} \begin{bmatrix} -0.4680 - 0.6328i \\ -0.1492 - 0.5839i \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0.4153 - 0.4248i \\ 0.7759 - 0.1713i \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} -0.4153 + 0.4248i \\ -0.7759 + 0.1713i \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0.4680 + 0.6328i \\ 0.1492 + 0.5839i \\ 0 \\ 0 \end{bmatrix} \\
 \mathbf{c}_4 &= \begin{Bmatrix} \begin{bmatrix} -0.3344 - 0.7316i \\ -0.1492 - 0.5839i \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} -0.5754 + 0.2224i \\ 0.7759 - 0.1713i \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0.5754 - 0.2224i \\ -0.7759 + 0.1713i \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0.3344 + 0.7316i \\ 0.1492 + 0.5839i \\ 0 \\ 0 \end{bmatrix} \\
 \mathbf{c}_5 &= \begin{Bmatrix} \begin{bmatrix} -0.1492 - 0.5839i \\ 0 \\ 0 \\ -0.4680 - 0.6328i \end{bmatrix} & \begin{bmatrix} 0.7759 - 0.1713i \\ 0 \\ 0 \\ 0.4153 - 0.4248i \end{bmatrix} & \begin{bmatrix} -0.7759 + 0.1713i \\ 0 \\ 0 \\ -0.4153 + 0.4248i \end{bmatrix} & \begin{bmatrix} 0.1492 + 0.5839i \\ 0 \\ 0 \\ 0.4680 + 0.6328i \end{bmatrix} \\
 \mathbf{c}_6 &= \begin{Bmatrix} \begin{bmatrix} -0.3344 - 0.7316i \\ -0.1492 - 0.5839i \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} -0.5754 + 0.2224i \\ 0.7759 - 0.1713i \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0.5754 - 0.2224i \\ -0.7759 + 0.1713i \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0.3344 + 0.7316i \\ 0.1492 + 0.5839i \\ 0 \\ 0 \end{bmatrix}
 \end{Bmatrix}
 \end{aligned}$$

The SER performance of the proposed codebooks along with those of the other existing ones are shown in Fig. 6. Observe that the proposed codebooks produce the best results. At $\text{SER}=10^{-5}$, we experience a coding gain of about 1.5 dB over the next best methods: “Zhang [6]” and “Ma [17]”. We also evaluate the SER performance of the proposed codebooks

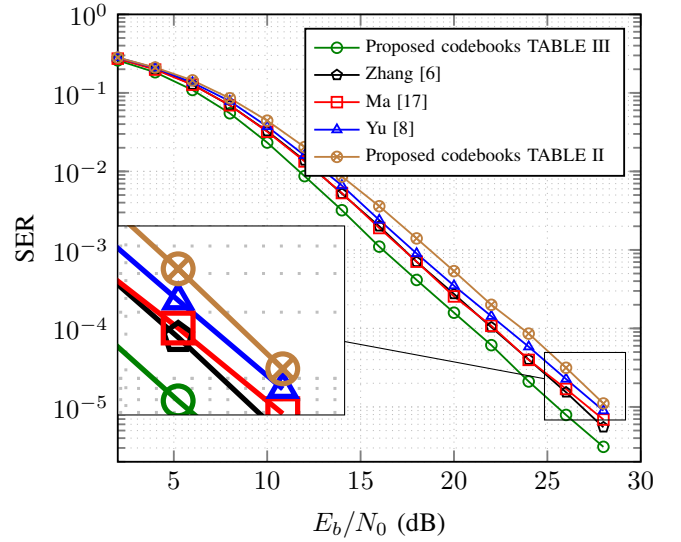


Fig. 6: SER performance of the SCMA system using various codebooks for $J = 6$ and $K = 4$ in Rayleigh fading channel.

designed for the AWGN channel. However, the performance is not satisfactory and inferior to “Zhang [6]”, “Ma [17]” and “Yu [8]”. This observation reiterates the well known fact that the optimum constellation for AWGN may not be optimum for Rayleigh fading channel and vice versa [13].

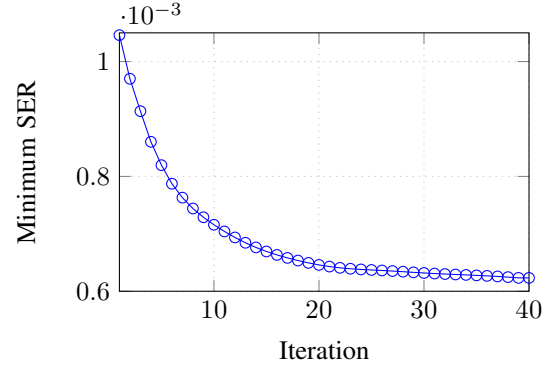


Fig. 7: Minimum SER versus iteration in Rayleigh fading channel at $\frac{E_b}{N_0} = 17$ dB during differential evolution process.

The progress of Algorithm 1 with increasing iteration is depicted in Fig. 7. The minimum SER corresponding to the best vector/row in the population matrix \mathbf{P} as updated in the current iteration is plotted against the iteration number. Observe from Fig. 7 that the minimum SER tends to settle down at a value after 20 iterations. We have observed similar progression of the DE-based algorithm in the case of AWGN channel. However, due to space constraint, the plot for AWGN channel is not included in the paper.

The values of the KPIs mentioned in Section II-A are shown in TABLE IV for various codebooks. “Zhang [6]” has the best Euclidean distance profile. Its $d_{E,\min}$ is the highest and τ_E is the lowest. Proposed codebook (AWGN) has the highest $d_{P,\min}$, however the τ_P is not the lowest. The Euclidean distance parameters for “Ma [17]” are the worst since $d_{E,\min}$ is the lowest and τ_E is the highest. Its poor performance over

AWGN channel can be attributed to this fact. However, its product distance profile is impressive. Its $d_{P,\min}$ is not the lowest and τ_P has the lowest value. However, it is difficult to justify the reported SER performances completely with the help of the KPIs mentioned in TABLE IV.

TABLE IV: Key performance indicators

	Proposed (AWGN)	Proposed (Fading)	Zhang [6]	Sharma [10]	Yu [8]	Ma [17]
$d_{E,\min}$	0.8966	0.8625	1.0171	0.9976	0.5351	0.3883
τ_E	4	4	2	2	2	4
$d_{P,\min}$	0.1103	0.0595	0.0810	0.0544	0.0379	0.0448
τ_P	4	4	4	4	2	2

These KPIs only partially characterize the SCMA system. The SCMA system is a complicated multi-user scenario where the detection is carried out by the sophisticated MPA. These KPIs fail to take the MPA-based detection process into account. Thus they are inadequate to facilitate a conclusive comparative analysis of various codebooks.

The above codebooks can be enlarged to build $J = 8$, $K = 4$ SCMA systems with 200% overloading factor. We consider the following factor matrix:

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

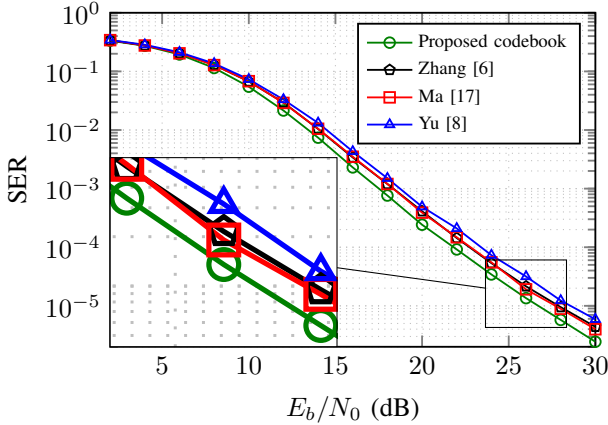


Fig. 8: SER performance of the SCMA system using various codebooks for $J = 8$ and $K = 4$ in Rayleigh fading channel.

The 3rd and the 4th columns are repeated as the 8th and 7th ones respectively. The constellation points for \mathcal{C}_3 and \mathcal{C}_4 are exchanged between the 7th and the 8th users. The SER performances of these codebooks are shown in Fig. 8 for Rayleigh fading channel. In this case also, the proposed DE-based codebooks yield the best result.

V. CONCLUSIONS

This paper presented a method to design the codebooks for an SCMA system. The constellation points are designed with the objective of minimizing the SER. The SER is considered as it directly reflects the effectiveness of the system to detect a user's data in interference-limited environment. First the structure of the codebooks is fixed using a finite number of complex

numbers. The minimization of the SER over these variables is accomplished with the help of DE. The optimum complex numbers are then used to form the desired codebooks. It is found that the codebook-design task is a channel-dependent affairs. The SER performance of the proposed codebooks for the AWGN and the fading channels are compared with those of other existing codebooks in literature. This comparison established the superiority of the proposed method over others.

REFERENCES

- [1] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74–81, 2015.
- [2] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1616–1626, April 2008.
- [3] H. Nikopour and H. Baligh, "Sparse code multiple access," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sept 2013, pp. 332–336.
- [4] F. R. Kschischang, B. J. Frey, and H. . Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, Feb 2001.
- [5] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, "SCMA codebook design," in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, Sep. 2014, pp. 1–5.
- [6] 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 *2016 8th International Conference on Wireless Communications Signal Processing (WCSP)*, Oct 2016, pp. 1–5.
- [7] M. Alam and Q. Zhang, "Designing optimum mother constellation and codebooks for SCMA," in *IEEE International Conference on Communications (ICC)*. IEEE, 2017, pp. 1–6.
- [8] L. Yu, X. Lei, P. Fan, and D. Chen, "An optimized design of SCMA codebook based on star-QAM signaling constellations," in *2015 International Conference on Wireless Communications Signal Processing (WCSP)*, Oct 2015, pp. 1–5.
- [9] J. Bao, Z. Ma, M. Xiao, T. A. Tsiftsis, and Z. Zhu, "Bit-interleaved coded SCMA with iterative multiuser detection: Multidimensional constellations design," *IEEE Transactions on Communications*, vol. 66, no. 11, pp. 5292–5304, Nov 2018.
- [10] S. Sharma, K. Deka, V. Bhatia, and A. Gupta, "SCMA codebook based on optimization of mutual information and shaping gain," in *2018 IEEE Globecom Workshops (GC Wkshps)*, Dec 2018, pp. 1–6.
- [11] R. Storn and K. Price, "Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, Dec 1997. [Online]. Available: <https://doi.org/10.1023/A:1008202821328>
- [12] S. Das and P. N. Suganthan, "Differential evolution: A survey of the state-of-the-art," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 1, pp. 4–31, Feb 2011.
- [13] J. Boutros and E. Viterbo, "Signal space diversity: a power- and bandwidth-efficient diversity technique for the rayleigh fading channel," *IEEE Transactions on Information Theory*, vol. 44, no. 4, pp. 1453–1467, July 1998.
- [14] M. Vameghestahbanati, I. D. Marsland, R. H. Gohary, and H. Yanikomeroglu, "Multidimensional constellations for uplink SCMA systems: a comparative study," *IEEE Communications Surveys Tutorials*, vol. 21, no. 3, pp. 2169–2194, thirdquarter 2019.
- [15] J. Denes and A. D. Keedwell, *Latin squares: New developments in the theory and applications*. Elsevier, 1991.
- [16] K. Xiao, B. Xia, Z. Chen, B. Xiao, D. Chen, and S. Ma, "On capacity-based codebook design and advanced decoding for sparse code multiple access systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 3834–3849, June 2018.
- [17] Mojtaba Vaezi, Zhiguo Ding and H. Vincent Poor, Ed., *Multiple Access Techniques for 5G Wireless Networks and Beyond*. Springer International, 2019.
- [18] K. Price, R. Storn, and J. Lampinen, *Differential Evolution*. Springer-Verlag Berlin Heidelberg, 2005.
- [19] Anyong Qing, *Differential Evolution: Fundamentals and Applications in Electrical Engineering*. Wiley IEEE, 2009.