

Uniquely Decodable Multi-Amplitude Sequence for Grant-Free Multiple-Access Adder Channels

Qi-Yue Yu, *Senior Member, IEEE*, and Ke-Xun Song

Abstract

Grant-free multiple-access (GFMA) is a valuable research topic, since it can support multiuser transmission with low latency. This paper constructs novel uniquely-decodable multi-amplitude sequence (UDAS) sets for GFMA systems, which can provide high spectrum efficiency (SE) with low-complexity active user detection (AUD) algorithm. First of all, we propose an UDAS-based multi-dimensional bit interleaving coded modulation (MD-BICM) transmitter; then introduce the definition of UDAS and construct two kinds of UDAS sets based on cyclic and quasi-cyclic matrix modes. Besides, we present a statistic of UDAS feature based AUD algorithm (SoF-AUD), and a joint multiuser detection and improved message passing algorithm for the proposed system. Finally, the active user error rate (AUER) and Shannon limits of the proposed system are deduced in details. Simulation results show that our proposed system can simultaneously support four users without additional redundancy, and the AUER can reach an extremely low value 10^{-5} when E_b/N_0 is 0 dB and the length of transmit block is larger than a given value, i.e., 784, verifying the validity and flexibility of the proposed UDAS sets.

Index Terms

Multiple access, pilot sequence, adder channel, Shannon limit, active user detection (AUD).

I. INTRODUCTION

Next generation multiple-access is expected to support massive users in the limited resources, and many works have been done on this topic [1]–[3]. In general, the multiple-access technique can be categorized into uncoordinated multiple-access and coordinated multiple-access [4]. The uncoordinated multiple-access is absent coordination and generally viewed as unsourced multiple-access, in which each user shares the same transmission protocol without allocated signature before transmission. Conversely, the coordinated scenario is to administer the users by a central processor, i.e., base station (BS), and each user is assigned some unique signature that can be recognized by the receiver for detection.

The grant-free multiple-access (GFMA) can be viewed as a special case of the coordinated multiple-access, where each user may access the BS randomly. The BS needs to detect both the number of active users and their corresponding data sequences. The major difference between a grant-free coordinated multiple-access and an uncoordinated multiple-access is relayed on the dedicated signature [4]. The grant-free scenario generally allocates a pilot sequence (or signature) to each user; thus, the receiver can separate and identify the users with the help of pilot sequences. In contrast, the active users of the uncoordinated multiple-access case randomly select pilot sequences without any coordination, and sometimes may lead to pilot collisions.

Therefore, it is important to design multiuser signatures for GFMA systems. The well-known pseudo random sequences are composed of binary bits $\{0, 1\}$, e.g., m-sequence, golden sequences [5], Reed-Muller codes [6], Walsh sequences, and etc. With the development of spreading sequences, many papers discuss uniquely-decodable (UD) ternary code sets $\{-1, 0, +1\}$ for the overloaded synchronous code division multiple-access (CDMA) systems [7]–[12], which can support larger number of users than the classical orthogonal spreading codes. Besides the binary and ternary code sequences, Zad-Off Chu (ZC) sequences are also popular pilot sequences, especially for channel estimation. It is found that the amplitude of ZC sequences is generally a constant, and only phases are varied [14]–[16]. In addition, some papers are interested in designing frameworks to avoid collisions for GFMA systems, e.g., paper [17] treats collisions as interference and builds the statistical model with the aid of Poisson point processes. Moreover, multiuser codebook design is also taken into consideration for GFMA. In [18], it presents multi-dimensional (MD) codebooks of sparse code multiple access (SCMA) in a grant-free multiple-access channel (MAC), where the data stream of each user is directly mapped to a codeword of

the proposed MD codebook. Due to the sparsity of MD codebooks, the proposed scheme can maintain overloaded information and enable massive connectivity. When these classical pilot sequences (or MD codebooks) are used as multiuser signatures for a GFMA system, there exist some challenges and/or drawbacks.

- The spectrum efficiency (SE) of spreading sequences based multiple-access systems is generally low. For example, when a binary spreading sequence (e.g. Walsh sequence) is utilized, the SE of each user is equal to the reciprocal of the spreading factor (SF), i.e., $\frac{1}{SF}$, and the sum-rate of the multiuser system is equal to or smaller than one. When UD ternary codes are utilized, the sum-rate can be larger than one, because of the overloaded information; however, the SE of each user is still equal to $\frac{1}{SF}$.
- The designed multiuser codebooks cannot be flexibly extended to a general multiple-access case. Most of the recent multiuser codebooks are designed based on multi-dimensional constellations (or lattice, and etc.), which have strict constraints on transmit signals' amplitudes and phases [19]–[23]. Thus, the extension of the designed multiuser codebooks is generally insignificant, especially for a massive random access system.
- Most of the recent active user detection (AUD) algorithms of GFMA systems are with relatively high complexities, and few of them focus on the theoretical analyses of the AUD processing. For example, the primary AUD methods are compressed sensing techniques [24]–[26], successive joint decoding (SJD) [17], successive interference cancellation (SIC), blind detections [18], and etc.

Regarding as the aforementioned challenges, it is interesting to design special sequences for GFMA systems. Until now, most of the classical pilot sequences are designed based on binary (or ternary, or phase) sets, few concerns on the multi-amplitude information. In [28], we have proposed the concept of uniquely-decodable mapping (UDM). It is declared that, if each user exploits 2ASK (amplitude shift keying) and the amplitudes of J users are respectively $\{1, 2^1, \dots, 2^{J-1}\}$, then the J users can be uniquely separated without ambiguity at the receiver. For example, if there are two users and the transmit signals of the two users are respectively $\{-1, +1\}$ and $\{-2, +2\}$, the superimposed signal set is $\{-3, -1, +1, +3\}$ that is a one-to-one mapping between the two users' transmit signals and the superimposed signals.

Based on the conception of UDM, this paper constructs a novel uniquely-decodable multi-amplitude sequence (UDAS) for GFMA systems. The advantages of the proposed UDAS include

three aspects. First of all, it presents a high SE without additional redundancy, likewise, some designed non-orthogonal multiple-access (NOMA) codebooks. Following, it can be easily applied to various multiple-access scenarios, with favoured flexible. Finally, the AUD algorithm can be easily realized. The major contributions of this paper are four-fold:

- 1) This paper firstly proposes an UDAS-based multi-dimensional bit interleaving coded modulation (MD-BICM) transmitter, which is a combination of two channel encoders, one interleaver and multi-dimensional modulation.
- 2) We define the conception of UDAS, and construct two kinds of UDAS sets based on cyclic and quasi-cyclic matrix modes. Some important features of the cyclic/quasi-cyclic UDAS sets are deduced in details.
- 3) We propose a statistic of UDAS feature based AUD algorithm (SoF-AUD), and a joint multiuser detection (MUD) and improved message passing iterative decoding algorithm for the proposed system.
- 4) Finally, we deduce the theoretical active user error rate (AUER) of the proposed system; then, analyze and calculate the Shannon limits of the multiple-access adder channels in details.

The remainder of this paper is organized as follows. In Section II, the UDAS-based MD-BICM system is described. We present some definitions of UDAS, and construct two kinds of UDAS sets with detailed features in Section III. Section IV proposes detection algorithms for the proposed system. The theoretical analyses are presented in Section V. Simulation results are discussed in Section VI, followed by concluding remarks drawn in Section VII.

II. SYSTEM MODEL

In this paper, a , \mathbf{a} and \mathbf{A} stand for a variable, a vector and a matrix, respectively. Denote \mathbf{A}^T by the transpose of a matrix \mathbf{A} . Let \mathbb{B} , \mathbb{Z} and \mathbb{C} be the binary, integer and complex fields, respectively. $\text{Re}[a]$ and $\text{Im}[a]$ are respectively the real part and image part of the complex number a . i stands for an image number. $E[\cdot]$ is the function of expectation, and $\lceil \cdot \rceil$ denotes the ceiling operator. $\|\mathbf{a}\|_2$ denote the 2-norm of a vector \mathbf{a} .

Suppose there are J users simultaneously access the BS. Obviously, the value of J should be estimated for a GFMA system. Each user is equipped with a single antenna, and the BS holds one antenna. The system model is shown in Fig. 1.

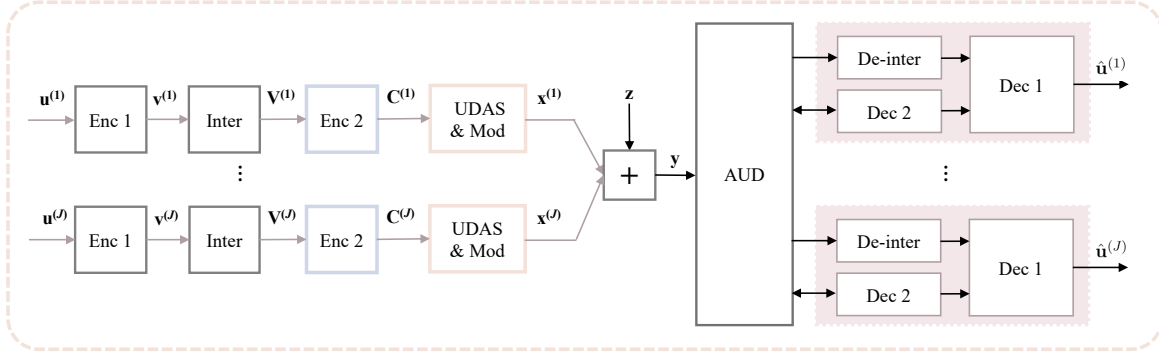


Fig. 1. A diagram of the proposed UDAS-based MD-BICM system, where “Enc”, “Dec”, “Inter”, “De-inter”, and “Mod” respectively stand for “encoder”, “decoder”, “interleaver”, “de-interleaver” and “modulation”.

The transmit bit information of the j th user is defined by $\mathbf{u}^{(j)} \in \mathbb{B}^{1 \times K}$, where $1 \leq j \leq J$ and K is the number of bits of the j th user. $\mathbf{u}^{(j)}$ is then passed to the first channel encoder whose generator is a $K \times N_1$ matrix defined by $\mathbf{G}_1 = [g_{k,n_1}]_{1 \leq k \leq K, 1 \leq n_1 \leq N_1}$, and obtained an encoded codeword $\mathbf{v}^{(j)} = \mathbf{u}^{(j)} \cdot \mathbf{G}_1$, where $\mathbf{v}^{(j)} = (v_1^{(j)}, v_2^{(j)}, \dots, v_{n_1}^{(j)}, \dots, v_{N_1}^{(j)})$. We can write $\mathbf{v}^{(j)}$ into a matrix $\mathbf{V}^{(j)}$ column-by-column, given as

$$\mathbf{V}^{(j)} = \begin{bmatrix} v_1^{(j)} & v_{M+1}^{(j)} & \cdots & v_{(N_c-1)M+1}^{(j)} \\ v_2^{(j)} & v_{M+2}^{(j)} & \cdots & v_{(N_c-1)M+2}^{(j)} \\ \vdots & \vdots & \ddots & \vdots \\ v_M^{(j)} & v_{2M}^{(j)} & \cdots & v_{N_c M}^{(j)} \end{bmatrix},$$

which is an $M \times N_c$ matrix with the condition $N_1 \leq M \cdot N_c$, and $v_{n_1}^{(j)} = 0$ for all $N_1 < n_1 \leq M \cdot N_c$. Actually, $\mathbf{V}^{(j)}$ is utilized to show the interleaving process, since the encoded codeword $\mathbf{v}^{(j)}$ will be further processed row-by-row, and the number of columns in $\mathbf{V}^{(j)}$ is directly related to the length of the selected UDAS.

Then, each row of $\mathbf{V}^{(j)}$ selects $L - 1$ bits, and passes the selected $L - 1$ bits to the second encoder \mathbf{G}_2 . In fact, \mathbf{G}_2 is designed for AUD, and it can be as simple as possible, e.g., single parity-check code (SPC). For the m th row of $\mathbf{V}^{(j)}$ for $1 \leq m \leq M$, let the parity-check bit generated by \mathbf{G}_2 be $v_{m,pc}^{(j)}$, which is located at the end of the m th row of $\mathbf{V}^{(j)}$. Thereafter, we can achieve an $M \times N$ matrix $\mathbf{C}^{(j)}$, as

$$\mathbf{C}^{(j)} = \begin{bmatrix} \mathbf{c}_1^{(j)} \\ \mathbf{c}_2^{(j)} \\ \vdots \\ \mathbf{c}_M^{(j)} \end{bmatrix} = \begin{bmatrix} v_1^{(j)} & v_{M+1}^{(j)} & \cdots & v_{(N_c-1)M+1}^{(j)} & v_{1,pc}^{(j)} \\ v_2^{(j)} & v_{M+2}^{(j)} & \cdots & v_{(N_c-1)M+2}^{(j)} & v_{2,pc}^{(j)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ v_M^{(j)} & v_{2M}^{(j)} & \cdots & v_{N_c M}^{(j)} & v_{M,pc}^{(j)} \end{bmatrix},$$

where $\mathbf{c}_m^{(j)} = (v_m^{(j)}, v_{M+m}^{(j)}, \dots, v_{(N_c-1)M+m}^{(j)}, v_{m,pc}^{(j)})$ is a $1 \times N$ vector, with $N = N_c + 1$.

Consequently, $\mathbf{C}^{(j)}$ is modulated row-by-row. Take the m th row $\mathbf{c}_m^{(j)}$ as example to explain the modulation mapping. Assume the multi-dimensional (MD) modulation index is \mathcal{M} and the length of an UDAS is L , under the assumption $N = L \cdot \log_2 \mathcal{M}$. For further discuss convenience, we rewrite $\mathbf{c}_m^{(j)}$ as

$$\begin{aligned} \mathbf{c}_m^{(j)} &= (\{c_{m,1,1}^{(j)}, c_{m,1,2}^{(j)}, \dots, c_{m,1,\log_2 \mathcal{M}}^{(j)}\}, \dots, \{c_{m,l,1}^{(j)}, c_{m,l,2}^{(j)}, \dots, c_{m,l,\log_2 \mathcal{M}}^{(j)}\}, \dots, \{c_{m,L,1}^{(j)}, c_{m,L,2}^{(j)}, \dots, c_{m,L,\log_2 \mathcal{M}}^{(j)}\}) \\ &= (\mathbf{c}_{m,1}^{(j)}, \dots, \mathbf{c}_{m,l}^{(j)}, \dots, \mathbf{c}_{m,L}^{(j)}) \end{aligned}$$

where $c_{m,l,b}^{(j)} \in \mathbb{B}$ for $1 \leq l \leq L$ and $1 \leq b \leq \log_2 \mathcal{M}$, and $\mathbf{c}_{m,l}^{(j)} = \{c_{m,l,1}^{(j)}, c_{m,l,2}^{(j)}, \dots, c_{m,l,\log_2 \mathcal{M}}^{(j)}\}$. Then, every $\log_2 \mathcal{M}$ bits of $\mathbf{c}_m^{(j)}$ are modulated to one symbol, i.e., $\mathbf{c}_{m,l}^{(j)} \rightarrow x_{m,l}^{(j)}$. We can achieve a modulated $M \times L$ matrix $\mathbf{X}^{(j)}$ as

$$\mathbf{X}^{(j)} = \begin{bmatrix} \mathbf{x}_1^{(j)} \\ \mathbf{x}_2^{(j)} \\ \dots \\ \mathbf{x}_M^{(j)} \end{bmatrix} = \begin{bmatrix} x_{1,1}^{(j)} & x_{1,2}^{(j)} & \dots & x_{1,L}^{(j)} \\ x_{2,1}^{(j)} & x_{2,2}^{(j)} & \dots & x_{2,L}^{(j)} \\ \vdots & \vdots & \ddots & \vdots \\ x_{M,1}^{(j)} & x_{M,2}^{(j)} & \dots & x_{M,L}^{(j)} \end{bmatrix},$$

where $x_{m,l}^{(j)}$ is an \mathcal{M}_1 -dimensional modulated symbol, where $\mathcal{M} = 2\mathcal{M}_1$, expressed as

$$(s_{m,l}^{(j)}, 0, 0, \dots, 0), (0, s_{m,l}^{(j)}, 0, \dots, 0), \dots, (0, 0, 0, \dots, s_{m,l}^{(j)}),$$

in which only one position of $x_{m,l}^{(j)}$ has value $s_{m,l}^{(j)}$ and the other $\mathcal{M}_1 - 1$ positions are all zeros, for $1 \leq m \leq M$ and $1 \leq l \leq L$. Note that \mathcal{M}_1 -dimensional modulation can increase both the flexibility and the minimum Euclidean distance, thus is appealing for GFMA. If the i th location of $x_{m,l}^{(j)}$ is non-zero, we set the location index $\rho_{m,l,i}^{(j)} = 1$, and other location indexes $\rho_{m,l,i'}^{(j)} = 0$ for $i' \neq i$ and $1 \leq i' \leq \mathcal{M}_1$. The non-zero location is determined by the first $\log_2 \mathcal{M}_1$ bits of $\mathbf{c}_{m,l}^{(j)}$.

Suppose the j th user utilizes $\mathbf{e}_j = (e_{j,1}, e_{j,2}, \dots, e_{j,L})$ as its UDAS, then $s_{m,l}^{(j)}$ is calculated as

$$s_{m,l}^{(j)} = \left(2c_{m,l,\log_2 \mathcal{M}}^{(j)} - 1\right) \cdot e_{j,l}, \quad (1)$$

which is determined by the last bit $c_{m,l,\log_2 \mathcal{M}}^{(j)}$ of $\mathbf{c}_{m,l}^{(j)}$ and the l th symbol of \mathbf{e}_j .

Look back to the second encoder \mathbf{G}_2 . Since \mathbf{G}_2 is used to assist AUD, the input $(L-1)$ bits of \mathbf{G}_2 are set to be the last bit of $\mathbf{c}_{m,l}^{(j)}$ for $1 \leq l \leq L$, then

$$\begin{aligned} v_{m,pc}^{(j)} &= c_{m,1,\log_2 \mathcal{M}}^{(j)} \oplus c_{m,2,\log_2 \mathcal{M}}^{(j)} \oplus \dots \oplus c_{m,L,\log_2 \mathcal{M}}^{(j)} \\ &= v_{m+M \cdot (\log_2 \mathcal{M}-1)}^{(j)} \oplus v_{m+M \cdot (2\log_2 \mathcal{M}-1)}^{(j)} \oplus \dots \oplus v_{m+M \cdot ((L-1)\log_2 \mathcal{M}-1)}^{(j)}. \end{aligned}$$

For better understand, an example is presented to show the interleaving and encoding process.

Example 1: Assume $\mathbf{v}^{(j)} = (0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 1, 0, 1)$, which is a 1×21 vector and $N_1 = 21$. Set $M = 2$, $N_c = 11$, $\mathcal{M} = 8$ and $\mathcal{M}_1 = 4$, then $\mathbf{v}^{(j)}$'s corresponding matrix $\mathbf{V}^{(j)}$ is given by

$$\mathbf{V}^{(j)} = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & \textcolor{red}{0} \end{bmatrix},$$

and $v_{2,11}$ is set to be 0 (shown in red color). Then, SPC is done to each row of $\mathbf{V}^{(j)}$, given by

$v_{1,p_1}^{(j)} = v_5 \oplus v_{11} \oplus v_{17} = 0$ and $v_{2,p_1}^{(j)} = v_6 \oplus v_{12} \oplus v_{18} = 1$. Thus, it is obtained

$$\mathbf{C}^{(j)} = \begin{bmatrix} \mathbf{c}_1^{(j)} \\ \mathbf{c}_2^{(j)} \end{bmatrix} = \begin{bmatrix} 0\textcolor{blue}{1}1 & 10\textcolor{blue}{0} & 00\textcolor{blue}{1} & 11\textcolor{red}{0} \\ 10\textcolor{blue}{0} & 01\textcolor{blue}{0} & 11\textcolor{blue}{1} & 00\textcolor{red}{1} \end{bmatrix},$$

which is a 2×12 matrix, where $N = 12$ and $L = 4$. End Example 1.

Then, $\mathbf{X}^{(j)}$ is sent to the adder multiple-access channel row-by-row. At the receiver, the received signal of the m th row \mathbf{y}_m is equal to

$$\mathbf{y}_m = \sum_{j=1}^J \mathbf{x}_m^{(j)} + \mathbf{z}_m, \quad (2)$$

where $1 \leq m \leq M$, and \mathbf{z}_m is an additive white Gaussian noise (AWGN) vector. Each noise component of \mathbf{z}_m is an independent and identically distributed (*i.i.d*) Gaussian random variable with distribution $\mathcal{N}(0, N_0/2)$. Set $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_M)$, where $\mathbf{y}_m = (y_{m,1}, y_{m,2}, \dots, y_{m,l}, \dots, y_{m,L})$ includes L symbols. For the l th received symbol $y_{m,l}$ of \mathbf{y}_m , we have

$$y_{m,l} = (y_{m,l,1}, y_{m,l,2}, \dots, y_{m,l,i}, \dots, y_{m,l,\mathcal{M}_1}),$$

which is an \mathcal{M}_1 -dimensional signal, and

$$y_{m,l,i} = \sum_{j=1}^J \rho_{m,l,i}^{(j)} s_{m,l}^{(j)} + z_{m,l,i}, \quad (3)$$

where $\rho_{m,l,i}^{(j)} = \{0, 1\}$. Based on the received \mathbf{y} , the AUD modular can detect the number of arrival users and identify users' signatures. Afterwards, by a joint detection algorithm, all the users' information bits can be recovered. It is noted that the two encoders (\mathbf{G}_1 and \mathbf{G}_2) can be removed, and the system still works, which reflects the great flexibility of the proposed scheme.

In the following discussion, the subscripts “ j ”, “ m ”, “ l ”, “ b ” and “ i ” still stand for the j th user, the m th row, the l th symbol, the b th bit of the symbol, and the i th location of the \mathcal{M}_1 -dimension modulation, where $1 \leq j \leq J, 1 \leq m \leq M, 1 \leq l \leq L, 1 \leq b \leq \log_2 \mathcal{M}$ and $1 \leq i \leq \mathcal{M}_1$.

III. DEFINITION, CONSTRUCTION, AND FEATURES OF UDAS SET

This section presents definition, construction and features of UDAS set in synchronous adder multiple-access channels. A good designed UDAS set can separate multiuser without ambiguous.

A. Definition

According to [28], an uniquely-decodable mapping (UDM) element set is defined by $\Delta = \{\Delta_{re}, \Delta_{im}\}$, where $\Delta_{re} = \{1, 2, \dots, 2^p\}$ and $\Delta_{im} = \{1i, 2i, \dots, 2^pi\}$ with $p \in \mathbb{Z}$. This UDM element set can maximum simultaneously support $2(p+1)$ users without ambiguous, if each user selects one of the element, i.e., a , and its inverse, i.e., $-a$, as the user's modulated symbols [28]. Denote $|\Delta|$, $|\Delta_{re}|$, and $|\Delta_{im}|$ by the number of elements in sets Δ , Δ_{re} , and Δ_{im} , respectively. Obviously, it is derived that $|\Delta_{re}| = |\Delta_{im}| = p+1$ and $|\Delta| = 2(p+1)$, which are determined by the parameter p .

Definition 1. An L -length UDAS is consisted of L sequential elements, which are selected from an UDM element set Δ with p as a parameter.

Let all the L -length UDAS belong to a space \mathbb{E}_L , and the total number of \mathbb{E}_L space is equal to $|\mathbb{E}_L| = |\Delta|^L = [2(p+1)]^L$.

Definition 2. Assume a set Ψ contains T sequences, as

$$\Psi = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_t, \dots, \mathbf{e}_T\},$$

with $\mathbf{e}_t = (e_{t,1}, e_{t,2}, \dots, e_{t,l}, \dots, e_{t,L}) \in \mathbb{E}_L$, where $e_{t,l} \in \Delta$ and $1 \leq l \leq L$. There are T users, and each user is assigned a sequence, i.e., \mathbf{e}_t , and the transmit vector of the t th user is defined by $\mathbf{c}^{(t)} = (c_1^{(t)}, c_2^{(t)}, \dots, c_l^{(t)}, \dots, c_L^{(t)})$ where $\mathbf{c}^{(t)} \in \mathbb{B}^{1 \times L}$ for $1 \leq t \leq T$. The modulated symbol of the l th symbol of the t th user is equal to $(2c_l^{(t)} - 1) \cdot e_{t,l}$, thus the sum-pattern of T users w_l can be given as $w_l = \sum_{t=1}^T (2c_l^{(t)} - 1) \cdot e_{t,l}$, where $w_l \in \Omega_l^T$ and $|\Omega_l^T| = 2^T$ for $1 \leq l \leq L$. Then, the sum-pattern vector of the entire L symbols of the T users is defined by $\mathbf{w} = \{w_1, w_2, \dots, w_l, \dots, w_L\}$ that belongs to the sum-pattern set Ω^T , i.e., $\mathbf{w} \in \Omega^T$, where $\Omega^T = \{\Omega_1^T, \Omega_2^T, \dots, \Omega_l^T, \dots, \Omega_L^T\}$. If the sequences in Ψ satisfy the following conditions:

- 1) When $t \neq t'$ and $1 \leq t, t' \leq T$, we have $e_{t,l} \neq e_{t',l}$ for all $1 \leq l \leq L$;
- 2) All the vectors in Ω^T are different from each other;
- 3) It is a one-to-one mapping between the sum-pattern vector and the transmit vectors of T users, i.e., $\mathbf{w} \leftrightarrow \{\mathbf{c}^{(1)}, \mathbf{c}^{(2)}, \dots, \mathbf{c}^{(t)}, \dots, \mathbf{c}^{(T)}\}$; and

- 4) *The power constraint satisfies $\lim_{L \rightarrow \infty} \mathbb{P}(|P_t - P_{avg}| < \varepsilon) = 0$, where $P_t = \frac{1}{L} \sum_{l=1}^L e_{t,l}^2$ stands for the average power of the t th UDAS, $P_{avg} = \frac{1}{T} \sum_{t=1}^T P_t$ is the average power of the entire UDAS set, and ε is a positive small number;*

then, the set Ψ is defined by a T -size UDAS set.

The former three conditions are used to ensure the superimposed signals of T users can be uniquely separated. The last condition anticipates the average power of each user keeps as a constant. For a multiple-access network, we hope the average transmit power of each user is almost the same, so that to keep fairness among users. Besides the aforementioned conditions, there are some supplementary conditions, for example, $\frac{\|\mathbf{e}_t\|_\infty}{\sqrt{P_{avg}}} < \epsilon$, which is used to make the peak-to-average power ratio (PAPR) of one user within an acceptable range.

B. Construction

There are many approaches to construct a T -size UDAS set. This paper presents two of them, which are cyclic and quasi-cyclic (QC) matrix based construction schemes.

1) *Cyclic matrix mode:* Let \mathbf{E}_{cyc} be a cyclic matrix, and the subscript “cyc” indicates “cyclic”. The first row of \mathbf{E}_{cyc} is viewed as a generator of \mathbf{E}_{cyc} , denoted by a $1 \times T$ vector, i.e., $\mathbf{a} = (a_1, a_2, \dots, a_t, \dots, a_T)$, where a_t is an element in Δ . Each element of \mathbf{a} is shifted to the rightward and downward one position, and the last element of \mathbf{a} is moved to the first position, achieving the second row of \mathbf{E}_{cyc} . Repeat this processing, until the first element a_1 of \mathbf{a} has researched the last position of the T th row. \mathbf{E}_{cyc} is a $T \times T$ matrix, given by

$$\mathbf{E}_{cyc} = \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_T \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \dots & a_T \\ a_T & a_1 & \dots & a_{T-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_2 & a_3 & \dots & a_1 \end{bmatrix}. \quad (4)$$

Because of the cyclic structure, the rows of \mathbf{E}_{cyc} , i.e., $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_T$, are formed a T -size UDAS set Ψ . Moreover, it is easy to derive that $P_{avg} = P_t = \frac{1}{L} \sum_{l=1}^L a_l^2$, where $1 \leq t \leq T$ and $L = T$. We give an example to show the construct processing.

Example 2: Assume $L = 4$, and the generator $\mathbf{a} = (1, 1i, 2, 2i)$. Then,

$$\mathbf{E}_{cyc} = \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \\ \mathbf{e}_4 \end{bmatrix} = \begin{bmatrix} 1 & 1i & 2 & 2i \\ 2i & 1 & 1i & 2 \\ 2 & 2i & 1 & 1i \\ 1i & 2 & 2i & 1 \end{bmatrix}, \quad (5)$$

The average power P_{avg} equals to 2.5. End Example 2.

Recall Example 1, we have obtained $\mathbf{c}_1^{(j)} = \{011 \ 100 \ 001 \ 110\}$ and $\mathbf{c}_2^{(j)} = \{100 \ 010 \ 111 \ 001\}$. Assume ‘00’, ‘01’, ‘10’ and ‘11’ are respectively mapped to four different locations, suppose $\mathbf{e}_j = (1, 1i, 2, 2i)$, it is able to obtain

$$\begin{aligned} x_{1,1}^{(j)} &= (0, +1, 0, 0), & x_{1,2}^{(j)} &= (0, 0, -1i, 0), & x_{1,3}^{(j)} &= (+2, 0, 0, 0), & x_{1,4}^{(j)} &= (0, 0, 0, -2i), \\ x_{2,1}^{(j)} &= (0, 0, -1, 0), & x_{2,2}^{(j)} &= (0, -1i, 0, 0), & x_{2,3}^{(j)} &= (0, 0, 0, +2), & x_{2,4}^{(j)} &= (+2i, 0, 0, 0), \end{aligned}$$

and the modulated $\mathbf{x}_1^{(j)}$ and $\mathbf{x}_2^{(j)}$ are transmitted to the channel.

2) *Quasi-cyclic matrix mode:* A T -size UDAS set with quasi-cyclic (QC) structure is defined by \mathbf{E}_{qc} , where the subscript “qc” stands for “quasi-cyclic”. Refer to the QC structure of QC-LDPC codes [29], let \mathbf{E}_{qc} be an $S \times Q$ array, and each position of \mathbf{E}_{qc} is corresponding to a matrix, as

$$\mathbf{E}_{qc} = \begin{bmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \dots & \mathbf{A}_{1,Q} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \dots & \mathbf{A}_{2,Q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{S,1} & \mathbf{A}_{S,2} & \dots & \mathbf{A}_{S,Q} \end{bmatrix}, \quad (6)$$

where $\mathbf{A}_{s,q}$ is an $\mathcal{L} \times \mathcal{L}$ cyclic matrix. Hence, \mathbf{E}_{qc} is an $(\mathcal{L} \cdot S) \times (\mathcal{L} \cdot Q)$ quasi-cyclic matrix. Denote the generator of $\mathbf{A}_{s,q}$ by $\mathbf{a}_{s,q}$, i.e., $\mathbf{a}_{s,q} = (a_{s,q,1}, a_{s,q,2}, \dots, a_{s,q,\mathcal{L}}, \dots, a_{s,q,\mathcal{L}})$, where $1 \leq s \leq S$, $1 \leq q \leq Q$ and $1 \leq \iota \leq \mathcal{L}$. Given q and ι , if $s \neq s'$, $\mathbf{a}_{s,q,\iota} \neq \mathbf{a}_{s',q,\iota}$, then the rows of \mathbf{E}_{qc} can form an $(\mathcal{L} \cdot S)$ -size UDAS set, in which each sequence is with length $L = \mathcal{L} \cdot Q$.

Equation (6) is a general expression of the QC structure, and it can be extended to a block-wise cyclic structure, given by

$$\mathbf{E}_{b,qc} = \begin{bmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \dots & \mathbf{A}_{1,Q} \\ \mathbf{A}_{1,Q} & \mathbf{A}_{1,1} & \dots & \mathbf{A}_{1,Q-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{1,2} & \mathbf{A}_{1,3} & \dots & \mathbf{A}_{1,1} \end{bmatrix}, \quad (7)$$

where $\mathbf{A}_{1,q}$ is an $\mathcal{L} \times \mathcal{L}$ cyclic matrix.

It is seen that the first row-block is shifted to the rightward and downward one position and the last block is moved to the first position, leading to a new row-block. Each row-block of (7) has cyclic structure, thus $\mathbf{E}_{b,qc}$ is an $(\mathcal{L} \cdot Q) \times (\mathcal{L} \cdot Q)$ block-wise cyclic square matrix, where “ b ” stands for “block-wise”.

Example 3: Let $Q = 3$, and the generators of $\mathbf{A}_{1,1}$, $\mathbf{A}_{1,2}$ and $\mathbf{A}_{1,3}$ are respectively $\mathbf{a}_{1,1} = (1, 1i)$, $\mathbf{a}_{1,2} = (2, 2i)$ and $\mathbf{a}_{1,3} = (4, 4i)$. Then, the block-wise QC structure $\mathbf{E}_{b,qc}$ is given by

$$\mathbf{E}_{b,qc} = \begin{bmatrix} 1 & 1i & 2 & 2i & 4 & 4i \\ 1i & 1 & 2i & 2 & 4i & 4 \\ 4 & 4i & 1 & 1i & 2 & 2i \\ 4i & 4 & 1i & 1 & 2i & 2 \\ 2 & 2i & 4 & 4i & 1 & 1i \\ 2i & 2 & 4i & 4 & 1i & 1 \end{bmatrix}. \quad (8)$$

Each row of $\mathbf{E}_{b,qc}$ is an UDAS, with average power $P_{avg} = 7$. End Example 3.

C. Features of the proposed cyclic/quasi-cyclic UDAS set

As aforementioned definition, the sum-pattern vector set of T -size UDAS set Ψ is $\Omega^T = \{\Omega_1^T, \Omega_2^T, \dots, \Omega_l^T, \dots, \Omega_L^T\}$. However, for a random multiple access scenario, not all the T users are simultaneously transmission. Now, we consider a general case.

Assume τ ($1 \leq \tau \leq T$) different sequences of Ψ are selected for multiuser transmission, i.e., $\Psi^{(\tau,\mu)} = \{\mathbf{e}_{t_1}, \dots, \mathbf{e}_{t_\nu}, \dots, \mathbf{e}_{t_\tau}\}$, where $t_\nu \in \{1, 2, \dots, T\}$ and μ is the τ -size combination index. Actually, there are totally C_T^τ combinations, indicating $1 \leq \mu \leq C_T^\tau$. At this moment, the sum-pattern of the l th symbol of the μ th combination index is $w_l = \sum_{\nu=1}^{\tau} (2c_l^{(\nu)} - 1) \cdot e_{t_\nu, l}$, where $w_l \in \omega_l^{(\tau,\mu)}$, and the number of elements in $\omega_l^{(\tau,\mu)}$ equals to $|\omega_l^{(\tau,\mu)}| = 2^\tau$.

Define the sum-pattern vector set of τ users by $\Omega^\tau = \{\Omega_1^\tau, \Omega_2^\tau, \dots, \Omega_l^\tau, \dots, \Omega_L^\tau\}$, where $\Omega_l^\tau = \omega_l^{(\tau,1)} \cup \omega_l^{(\tau,2)} \cup \dots \cup \omega_l^{(\tau,\mu)} \dots \cup \omega_l^{(\tau,C_T^\tau)}$ and the number of elements in Ω_l^τ satisfies $|\Omega_l^\tau| \leq C_T^\tau \cdot 2^\tau$. When $\tau = T$, it is found that $|\Omega_l^T| = 2^T$, equaling to the T -size case.

Moreover, define the maximum values of the real and image parts of $\omega_l^{(\tau,\mu)}$ by

$$\kappa_{l,re}^{(\tau,\mu)} = \max \left\{ \text{Re}[\omega_l^{(\tau,\mu)}] \right\}, \kappa_{l,im}^{(\tau,\mu)} = \max \left\{ \text{Im}[\omega_l^{(\tau,\mu)}] \right\}.$$

While, $(\kappa_{l,re}^{(\tau,\mu)}, \kappa_{l,im}^{(\tau,\mu)})$ are used for finding $\omega_l^{(\tau,\mu)}$. The average power of the sum-patterns in $\omega_l^{(\tau,\mu)}$, defined by $\lambda_l^{(\tau,\mu)}$, is equal to

$$\lambda_l^{(\tau,\mu)} = \frac{1}{2^\tau} \sum_{w_l \in \omega_l^{(\tau,\mu)}} \|w_l\|^2 = \sum_{\nu=1}^{\tau} e_{t_\nu, l}^2,$$

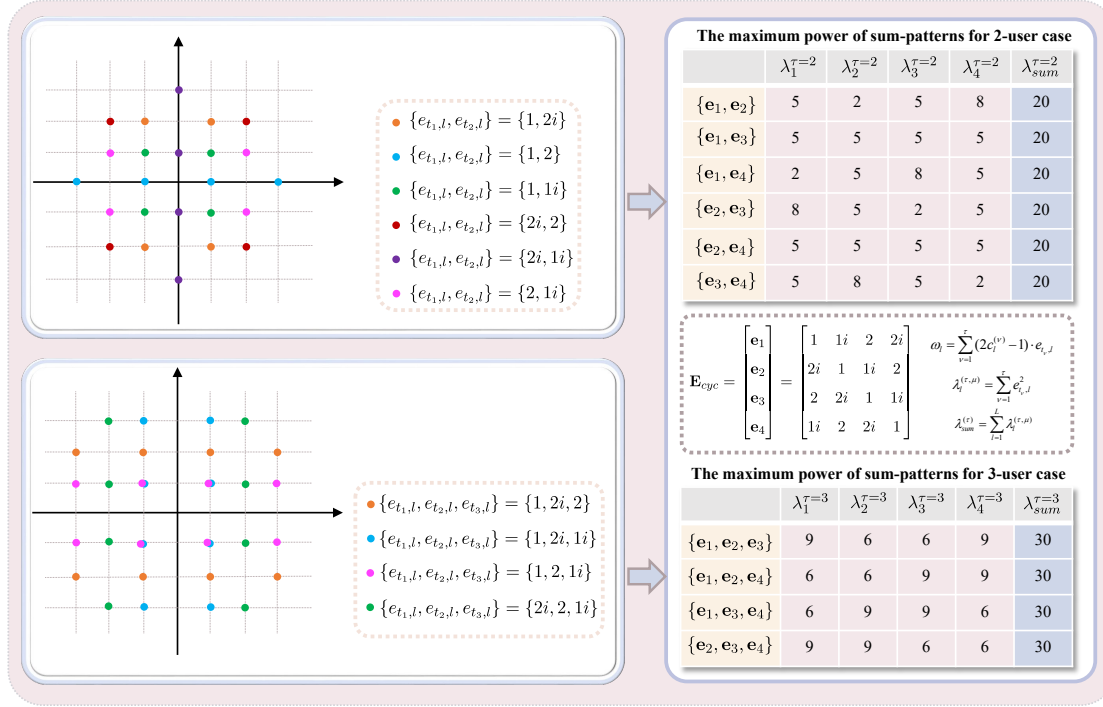


Fig. 2. An illustration of Example 4, where $\Psi = \{e_1, e_2, e_3, e_4\} = \{(1, 1i, 2, 2i), (2i, 1, 1i, 2), (2, 2i, 1, 1i), (1i, 2, 2i, 1)\}$. The upper part of the figure is for the case of $\tau = 2$, and the lower part of the figure is for the case of $\tau = 3$.

relying on the values of $\{e_{t_1,l}, e_{t_2,l}, \dots, e_{t_\nu,l}, \dots, e_{t_\tau,l}\}$. For a given τ , define the average power of $\omega_l^{(\tau,\mu)}$ of the total L locations by $\Lambda^{(\tau,\mu)} = \{\lambda_1^{(\tau,\mu)}, \lambda_2^{(\tau,\mu)}, \dots, \lambda_l^{(\tau,\mu)}, \dots, \lambda_L^{(\tau,\mu)}\}$, assisting AUD and multiuser detection. It is easy to derive that the sum of $\lambda_l^{(\tau,\mu)}$ is

$$\lambda_{sum}^{\tau} = \sum_{l=1}^L \lambda_l^{(\tau,\mu)} = \sum_{\nu=1}^{\tau} \sum_{l=1}^L e_{t_\nu,l}^2 = \tau \cdot LP_{avg},$$

which is a constant for a given τ , independent of the selection of sequences, i.e., combination index μ . This interesting result can help us quickly detect the number of arrival users. To illustrate this result, let us have a look at another example.

Example 4: Recall (5) of Example 2, there are four sequences, $\Psi = \{e_1, e_2, e_3, e_4\} = \{(1, 1i, 2, 2i), (2i, 1, 1i, 2), (2, 2i, 1, 1i), (1i, 2, 2i, 1)\}$, with $P_{avg} = 2.5$ and $L \cdot P_{avg} = 10$.

If we select any $\tau = 2$ sequences from the given Ψ for multiuser transmission, there are totally $C_4^2 = 6$ combinations, i.e., $\Psi^{(2,1)} = \{e_1, e_2\}$, $\Psi^{(2,2)} = \{e_1, e_3\}$, $\Psi^{(2,3)} = \{e_1, e_4\}$, $\Psi^{(2,4)} = \{e_2, e_3\}$, $\Psi^{(2,5)} = \{e_2, e_4\}$, and $\Psi^{(2,6)} = \{e_3, e_4\}$. The size of the sum-pattern set $\omega_l^{(\tau,\mu)}$ is

$|\omega_l^{(\tau,\mu)}| = 2^2 = 4$, where $1 \leq \mu \leq C_4^2$. When $l = 1$, we have

$$\begin{aligned} \{1, 2i\} &\rightarrow \omega_1^{(2,1)} = \{1 + 2i, 1 - 2i, -1 + 2i, -1 - 2i\} \rightarrow \lambda_1^{(2,1)} = 5; \\ \{1, 2\} &\rightarrow \omega_1^{(2,2)} = \{-3, -1, 1, 3\} \rightarrow \lambda_1^{(2,2)} = 5; \\ \{1, 1i\} &\rightarrow \omega_1^{(2,3)} = \{1 + 1i, 1 - 1i, -1 + 1i, -1 - 1i\} \rightarrow \lambda_1^{(2,3)} = 2; \\ \{2i, 2\} &\rightarrow \omega_1^{(2,4)} = \{2 + 2i, 2 - 2i, -2 + 2i, -2 - 2i\} \rightarrow \lambda_1^{(2,4)} = 8; \\ \{2i, 1i\} &\rightarrow \omega_1^{(2,5)} = \{-3i, -1i, 1i, 3i\} \rightarrow \lambda_1^{(2,5)} = 5; \\ \{2, 1i\} &\rightarrow \omega_1^{(2,6)} = \{2 + 1i, 2 - 1i, -2 + 1i, -2 - 1i\} \rightarrow \lambda_1^{(2,6)} = 5. \end{aligned}$$

It is found that $\omega_l^{(2,\mu)} \cap \omega_l^{(2,\mu')} = \emptyset$, when $\mu \neq \mu'$. Moreover, it is able to derive that $\lambda_{sum}^{\tau=2} = 20$, which equals to $2 \times (L \cdot P_{avg})$.

Observe $\Lambda^{(2,\mu)}$, it is found that $\Lambda^{(2,2)} = \Lambda^{(2,5)}$, i.e., $(5, 5, 5, 5)$; thus, $\Psi^{(2,2)}$ and $\Psi^{(2,5)}$ cannot be separated based on their Λ . However, observe $\omega_1^{(2,2)}$ and $\omega_1^{(2,5)}$ again, it is found that $(\kappa_{l,re}^{(2,2)}, \kappa_{l,im}^{(2,2)}) = (3, 0)$ and $(\kappa_{l,re}^{(2,5)}, \kappa_{l,im}^{(2,5)}) = (0, 3)$, indicating that $\Psi^{(2,2)}$ and $\Psi^{(2,5)}$ can be separated based on their corresponding $(\kappa_{l,re}^{(\tau,\mu)}, \kappa_{l,im}^{(\tau,\mu)})$.

If we pick any $\tau = 3$ sequences from Ψ for multiuser transmission, there are $C_4^3 = 4$ combinations, i.e., $\Psi^{(3,1)} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, $\Psi^{(3,2)} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4\}$, $\Psi^{(3,3)} = \{\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_4\}$, and $\Psi^{(3,4)} = \{\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$. At this time, the sum-pattern sets of the μ th combination of the l th symbol can be given. Take $l = 1$ as an example, they are

$$\begin{aligned} \{1, 2i, 2\} &\rightarrow \omega_1^{(3,1)} = \{-3 + 2i, -3 - 2i, -1 + 2i, -1 - 2i, 1 + 2i, 1 - 2i, 3 + 2i, 3 - 2i\} \rightarrow \lambda_1^{(3,1)} = 9; \\ \{1, 2i, 1i\} &\rightarrow \omega_1^{(3,2)} = \{1 - 3i, -1 - 3i, 1 - 1i, -1 - 1i, 1 + 1i, -1 + 1i, 1 + 3i, -1 - 3i\} \rightarrow \lambda_1^{(3,2)} = 6; \\ \{1, 2, 1i\} &\rightarrow \omega_1^{(3,3)} = \{-3 + 1i, -3 - 1i, -1 + 1i, -1 - 1i, 1 + 1i, 1 - 1i, 3 + 1i, 3 - 1i\} \rightarrow \lambda_1^{(3,3)} = 6; \\ \{2i, 2, 1i\} &\rightarrow \omega_1^{(3,4)} = \{2 - 3i, -2 - 3i, 2 - 1i, -2 - 1i, 2 + 1i, -2 + 1i, 2 + 3i, -2 + 3i\} \rightarrow \lambda_1^{(3,4)} = 9. \end{aligned}$$

For this scenario, we can know $\lambda_{sum}^{\tau=3} = 30$ that is equal to $3 \times (L \cdot P_{avg})$. Moreover, it is found that $\omega_l^{(3,2)} \cap \omega_l^{(3,3)} = \{1 - 1i, -1 - 1i, -1 + 1i, -1 - 1i\}$, indicating that there exists constellation overlap. Observe the four sets $\Lambda^{(3,1)}, \Lambda^{(3,2)}, \Lambda^{(3,3)}$ and $\Lambda^{(3,4)}$, they are different from each other, thus it is a one-to-one mapping between $\Lambda^{(3,\mu)}$ and $\Psi^{(3,\mu)}$, i.e., $\Lambda^{(\tau,\mu)} \rightarrow \Psi^{(\tau,\mu)}$. For this scenario, we can calculate $\Lambda^{(\tau,\mu)}$, and then obtain the corresponding $\Psi^{(\tau,\mu)}$. End Example 4.

IV. ACTIVE USER DETECTION AND MULTIUSER DETECTION

This section presents a statistic of UDAS feature based active user detection (SoF-AUD), and a joint multiuser detection (MUD) and improved iterative message passing algorithm (MPA) for the proposed system. Assume that the users are assigned different sequences from the T -size

UDAS set Ψ . The receiver will detect the number of arrival users and recover their individual data information.

A. MAP detection

First of all, we introduce the maximum *a posteriori* (MAP) detection algorithm, which is an optimal solution for the proposed system. Take (1) into (3), $y_{m,l,i}$ is rewritten as

$$y_{m,l,i} = \sum_{j=1}^J \rho_{m,l,i}^{(j)} \cdot \left(2c_{m,l,\log_2 \mathcal{M}}^{(j)} - 1 \right) \cdot e_{j,l} + z_{m,l,i} = w_{m,l,i} + z_{m,l,i}, \quad (9)$$

where $w_{m,l,i} = \sum_{j=1}^J \rho_{m,l,i}^{(j)} \cdot \left(2c_{m,l,\log_2 \mathcal{M}}^{(j)} - 1 \right) \cdot e_{j,l}$ is the received superimposed signal (or called sum-pattern). For further discussion, set $\mathbf{w} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m, \dots, \mathbf{w}_M)$, and $\mathbf{w}_m = (\mathbf{w}_{m,1}, \mathbf{w}_{m,2}, \dots, \mathbf{w}_{m,l}, \dots, \mathbf{w}_{m,L})$, where $\mathbf{w}_{m,l} = (w_{m,l,1}, w_{m,l,2}, \dots, w_{m,l,i}, \dots, w_{m,l,\mathcal{M}_1})$.

Equation (9) includes variables J , $\rho_{m,l,i}^{(j)}$, $c_{m,l,\log_2 \mathcal{M}}^{(j)}$ and $e_{j,l}$. Actually, $\rho_{m,l,i}^{(j)}$ and $c_{m,l,\log_2 \mathcal{M}}^{(j)}$ together are corresponding to $c_{m,l}^{(j)}$. Based on the MAP criterion, it is able to derive that

$$(J, \mathbf{C}, \Psi^{(J,\mu)}) = \operatorname{argmax} \left[P_J \cdot \exp \left\{ -\frac{\|\mathbf{y} - \mathbf{w}\|^2}{N_0} \right\} \right], \quad (10)$$

where $\mathbf{C} = \{\mathbf{C}^{(1)}, \dots, \mathbf{C}^{(J)}\}$ is the set of transmit bits of all J users, $\Psi^{(J,\mu)}$ is the selected UDAS set, and P_J is probability of J -user simultaneously access the receiver. Assume the receiver can maximum detect J_{max} users. When $J > J_{max}$, the detection is interrupted and all the users' packets are lost. Considering the application of the T -size UDAS set Ψ , in the following discussion, we assume $J_{max} = T$ and $J = \tau$.

Although the MAP detection is the optimal algorithm, the complexity is extremely high for the proposed system, i.e., $\mathcal{O}\left((L \cdot \mathcal{M}^{ML})^T\right)$. Thereby, regarding as the features of the cyclic/quasi-cyclic UDAS set, we firstly introduce a SoF-AUD algorithm, whose complexity is extremely low.

B. Step 1: SoF-AUD

The SoF-AUD includes two parts, one is to detect the number of arrival users τ ($1 \leq \tau \leq T$), and the other is to find the selected UDAS set $\Psi^{(\tau,\mu)}$.

For the l th received symbol of \mathbf{y}_m , define the i th location of the l th received symbol of all M rows by $\mathbf{y}_{l,i}^{\text{loc}} = (y_{1,l,i}, y_{2,l,i}, \dots, y_{m,l,i}, \dots, y_{M,l,i})$. If we ignore the effect of noise and M is

large enough, $\mathbf{y}_{l,i}^{\text{loc}}$ can traverse all the constellations of $\omega_l^{(\tau,\mu)}$, including $(\kappa_{l,re}^{(\tau,\mu)}, \kappa_{l,im}^{(\tau,\mu)})$. Actually, the proposed SoF-AUD algorithm is relayed on the statistical properties of $\mathbf{y}_{l,i}^{\text{loc}}$.

The sum power γ_{sum} of the entire L symbols can be calculated as

$$\gamma_{\text{sum}} = \sum_{l=1}^L \gamma_l = \sum_{l=1}^L \left\{ \frac{1}{M} \sum_{m=1}^M \sum_{i=1}^{\mathcal{M}_1} \|y_{m,l,i}\|^2 \right\}, \quad (11)$$

where $\gamma_l = \frac{1}{M} \sum_{m=1}^M \left[\sum_{i=1}^{\mathcal{M}_1} \|y_{m,l,i}\|^2 \right]$ is the average power of the l th symbol.

Thus, for a given number of arrival users τ , the statistical mean of γ_{sum} can be deduced as

$$\begin{aligned} \bar{\gamma}_{\text{sum}}^\tau &= \mathbb{E} \left[\sum_{l=1}^L \left\{ \frac{1}{M} \sum_{m=1}^M \sum_{i=1}^{\mathcal{M}_1} \|y_{m,l,i}\|^2 \right\} \right] = \mathbb{E} \left[\sum_{l=1}^L \sum_{i=1}^{\mathcal{M}_1} \|w_{m,l,i} + z_{m,l,i}\|^2 \right], \\ &= \sum_{l=1}^L \sum_{i=1}^{\mathcal{M}_1} \mathbb{E} [\|w_{m,l,i} + z_{m,l,i}\|^2] = \lambda_{\text{sum}}^\tau + L\mathcal{M}_1 \cdot N_0 \end{aligned} \quad (12)$$

since $\mathbb{E}[w_{m,l,i} \cdot z_{m,l,i}] = 0$ and $\mathbb{E}[\|z_{m,l,i}\|^2] = N_0$. Equation (12) is the sum of $\lambda_{\text{sum}}^\tau$ and $L\mathcal{M}_1 \cdot N_0$, indicating $\bar{\gamma}_{\text{sum}}^\tau$ is a constant for a given τ . Here, $\lambda_{\text{sum}}^\tau = \tau \cdot (LP_{\text{avg}})$, and LP_{avg} is a constant for a given cyclic/quasi-cyclic UDAS set Ψ .

Since we utilize \mathcal{M}_1 -dimensional modulation, the constellation with the maximum power of the l th symbol in $\mathbf{y}_{l,i}^{\text{loc}}$ can be calculated as

$$\zeta_{l,re} = \sum_{i=1}^{\mathcal{M}_1} \left[\max_{1 \leq m \leq M} \{\text{Re}[\mathbf{y}_{l,i}^{\text{loc}}]\} \right], \quad \zeta_{l,im} = \sum_{i=1}^{\mathcal{M}_1} \left[\max_{1 \leq m \leq M} \{\text{Im}[\mathbf{y}_{l,i}^{\text{loc}}]\} \right].$$

At this time, we can estimate the number of arrival users based on minimum square error (MSE) criterion as

$$\begin{aligned} (\hat{\tau}, \tilde{\mu}) &= \arg \min_{1 \leq \tau \leq T} \|\gamma_{\text{sum}} - \bar{\gamma}_{\text{sum}}^\tau\| = \arg \min_{1 \leq \tau \leq T} \|\gamma_{\text{sum}} - (\lambda_{\text{sum}}^\tau + L\mathcal{M}_1 \cdot N_0)\|, \\ \text{s.t. } C1 : \hat{\mu} &= \arg \min_{\mu' = \tilde{\mu}} \sum_{l=1}^L d_l^{(\tau, \mu')}, \end{aligned} \quad (13)$$

where $d_l^{(\tau, \mu')} = \|\zeta_{l,re} - \kappa_{l,re}^{(\tau, \mu')}\|_2 + \|\zeta_{l,im} - \kappa_{l,im}^{(\tau, \mu')}\|_2$. Consider there are more than one $\tilde{\mu}$ can satisfy (13), $C1$ is used to find the only one UDAS set.

According to (13), it is able to obtain the number of arrival users $\hat{\tau}$ and the selected UDAS set $\Psi^{(\hat{\tau}, \hat{\mu})}$. The complexity of the AUD modular is $\mathcal{O}(ML^2)$.

C. Step 2: MUD

Based on the $\hat{\tau}$ and $\Psi^{(\hat{\tau}, \hat{\mu})}$ obtained by SoF-AUD, we begin to detect and separate the superimposed signals. Actually, we have taken into consideration the connection (or statistics) among the received signals of M rows (i.e., $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_M$) during the AUD process. Now, we can deal with the received signals row-by-row to reduce the complexity.

1) $\mathcal{M}_1 = 1$: When $\mathcal{M}_1 = 1$, $x_{m,l}^{(j)}$ is equivalent to $s_{m,l}^{(j)}$. The detected results $\hat{w}_{m,l,1}$ is selected from the set $\Psi^{(\hat{\tau},\hat{\mu})}$, expressed as

$$\begin{aligned} \hat{w}_{m,l,1} &= \arg \min_{\omega_{m,l,1} \in \Psi^{(\hat{\tau},\hat{\mu})}} \|y_{m,l,1} - \omega_{m,l,1}\|_2, \\ s.t. \quad C2 : \Psi^{(\hat{\tau},\hat{\mu})} &\rightarrow \hat{w}_{m,l,1} \rightarrow \{c_{m,l,1}^{(1)}, c_{m,l,1}^{(2)}, \dots, c_{m,l,1}^{(\hat{\tau})}\}, \\ C3 : 0 &= c_{m,l,1}^{(j)} \oplus c_{m,2,1}^{(j)} \oplus \dots \oplus c_{m,L,1}^{(j)}. \end{aligned} \quad (14)$$

$C2$ stands for the one-to-one mapping between $\hat{w}_{m,l,1}$ and $\{c_{m,l,1}^{(1)}, c_{m,l,1}^{(2)}, \dots, c_{m,l,1}^{(\hat{\tau})}\}$. $C3$ reflects the SPC constraint, where $1 \leq j \leq \hat{\tau}$.

It is remarkable that (14) is a hard decision processing, which helps fast realize MUD with acceptable BER. It can be applied to the scenario without \mathbf{G}_1 .

2) $\mathcal{M}_1 > 1$: When $\mathcal{M}_1 > 1$, the detection processing becomes a little complex. Define the sum-pattern set of the i th location of the \mathcal{M}_1 -dimensional modulation by $\Psi^{(\hat{\tau}_i, \hat{\mu}_i)}$, where $\hat{\tau}_i$ and $\hat{\mu}_i$ are respectively $\hat{\tau}_i$ users and the $\hat{\mu}_i$ th combination at the i th location. Let the set of \mathcal{M}_1 -dimensional sum-pattern be $\Theta^{(\hat{\tau}, \hat{\mu})} = (\Psi^{(\hat{\tau}_1, \hat{\mu}_1)}, \Psi^{(\hat{\tau}_2, \hat{\mu}_2)}, \dots, \Psi^{(\hat{\tau}_i, \hat{\mu}_i)}, \dots, \Psi^{(\hat{\tau}_{\mathcal{M}_1}, \hat{\mu}_{\mathcal{M}_1})})$, the detected results are then expressed as

$$\begin{aligned} \hat{w}_{m,l} &= \arg \min_{\theta_{m,l} \in \Theta^{(\hat{\tau}, \hat{\mu})}} \|y_{m,l} - \theta_{m,l}\|_2, \\ s.t. \quad C4 : \hat{\tau} &= \sum_{i=1}^{\mathcal{M}_1} \hat{\tau}_i, \\ C5 : \mu_i &\subseteq \mu, \\ C6 : \hat{w}_{m,l} &\rightarrow \{\mathbf{c}_{m,l}^{(1)}, \mathbf{c}_{m,l}^{(2)}, \dots, \mathbf{c}_{m,l}^{(\hat{\tau})}\}, \\ C7 : 0 &= c_{m,1,\log_2 \mathcal{M}}^{(j)} \oplus c_{m,2,\log_2 \mathcal{M}}^{(j)} \oplus \dots \oplus c_{m,L,\log_2 \mathcal{M}}^{(j)}. \end{aligned} \quad (15)$$

where $\hat{w}_{m,l} = (\hat{w}_{m,l,1}, \hat{w}_{m,l,2}, \dots, \hat{w}_{m,l,\mathcal{M}_1})$, $\theta_{m,l} = (\omega_{m,l,1}, \omega_{m,l,2}, \dots, \omega_{m,l,\mathcal{M}_1})$ and $\omega_{m,l,i} \in \Psi^{(\hat{\tau}_i, \hat{\mu}_i)}$ for $1 \leq i \leq \mathcal{M}_1$. $C4$ is to keep the sum of the detected number of users be a constant. $C5$ indicates that the users are from the same selected UDAS set. $C6$ shows the one-to-one mapping between $\hat{w}_{m,l}$ and $\{\mathbf{c}_{m,l}^{(1)}, \mathbf{c}_{m,l}^{(2)}, \dots, \mathbf{c}_{m,l}^{(\hat{\tau})}\}$. $C7$ also stands for the SPC constraint of \mathbf{G}_2 , where $1 \leq j \leq \hat{\tau}$.

Actually, (15) can be viewed as an extension of (14) to multi-dimensional case. The mainly difference is that the referred set is from $\Psi^{(\tau, \mu)}$ to $\Theta^{(\tau, \mu)}$, with complexity $\mathcal{O}(\mathcal{M}^{ML\hat{\tau}})$.

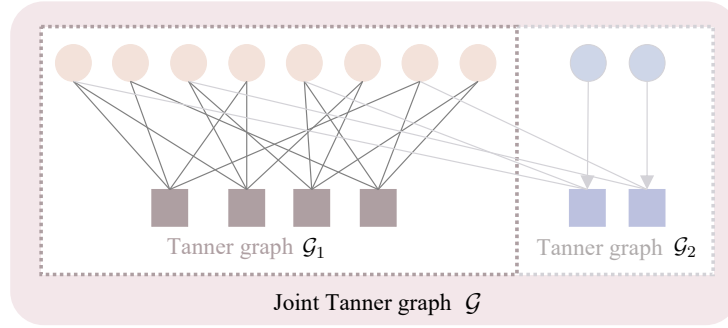


Fig. 3. A joint Tanner graph of the two encoders \mathcal{G} , where \mathcal{G}_1 and \mathcal{G}_2 are Tanner graphs of encoder one and encoder two, respectively.

D. Step 3: Improved MPA

When a LDPC code is utilized as the first encoder \mathbf{G}_1 , it is important to find the initial LLR (log likelihood ratio) according to the received signals for further message passing decoding. In other words, we should calculate the LLRs from the MUD process.

1) *Initial LLRs*: The probability of $\hat{\omega}_{m,l}$ can be calculated as

$$P(\hat{\omega}_{m,l}) = \frac{1}{\sqrt{\pi N_0}} \exp \left\{ -\frac{\|y_{m,l} - \hat{\omega}_{m,l}\|^2}{N_0} \right\}, \quad (16)$$

where $\hat{\omega}_{m,l} \in \Theta(\hat{\tau}, \hat{\mu})$. Since $\hat{\omega}_{m,l}$ is corresponding to $\log_2 \mathcal{M}$ bits of $\hat{\tau}$ users, i.e.,

$$\hat{\omega}_{m,l} \rightarrow \{\mathbf{c}_{m,l}^{(1)}, \mathbf{c}_{m,l}^{(2)}, \dots, \mathbf{c}_{m,l}^{(\hat{\tau})}\} = \{(c_{m,l,1}^{(1)}, c_{m,l,2}^{(1)}, \dots, c_{m,l,\log_2 \mathcal{M}}^{(1)}), (c_{m,l,1}^{(2)}, c_{m,l,2}^{(2)}, \dots, c_{m,l,\log_2 \mathcal{M}}^{(2)}), \dots, (c_{m,l,1}^{(\hat{\tau})}, c_{m,l,2}^{(\hat{\tau})}, \dots, c_{m,l,\log_2 \mathcal{M}}^{(\hat{\tau})})\},$$

the LLR of each bit, i.e., $c_{m,l,b}^{(j)}$ for $1 \leq b \leq \log_2 \mathcal{M}$, is calculated based on (16). We have

$$P(c_{m,l,b}^{(j)}) = \sum_{\hat{\omega}_{m,l,i} \rightarrow c_{m,l,b}^{(j)}} P(\hat{\omega}_{m,l,i}), \quad (17)$$

indicating that the probability of the b th bit of $\mathbf{c}_{m,l}^{(j)}$ is decided by the mapping relationship among $\mathbf{c}_{m,l}^{(j)}$, $x_{m,l}^{(j)}$, and $\hat{\omega}_{m,l}$, e.g., $\mathbf{c}_{m,l}^{(j)} \rightarrow x_{m,l}^{(j)} \rightarrow \hat{\omega}_{m,l}$ and vice versa.

Then, we can obtain the initial LLR of $c_{m,l,b}^{(j)}$ as

$$\text{LLR}(c_{m,l,b}^{(j)}) = \text{Log} \left[\frac{P(c_{m,l,b}^{(j)} = 0)}{P(c_{m,l,b}^{(j)} = 1)} \right], \quad (18)$$

where $1 \leq b \leq \log_2 \mathcal{M}$, $1 \leq l \leq L$ and $1 \leq m \leq M$. Thereafter, the LLRs of $\mathbf{C}^{(j)}$ are obtained. By de-interleaving $\mathbf{C}^{(j)}$ column-by-column, we can achieve the soft LLRs of $\mathbf{v}^{(j)}$ that are used for further decoding.

2) *Improved MPA*: Assume the parity-check matrix of the first encoder is \mathbf{H}_1 . Since the second encoder utilizes SPC, encoders one and two can be jointly decoding.

Let the Tanner graph of the two encoders be \mathcal{G}_1 and \mathcal{G}_2 , and the two graphs are connected by some of the variable nodes (VNs) of \mathcal{G}_1 and the check nodes (CNs) of \mathcal{G}_2 . Thus, the VNs update and CNs update should take into consideration the joint Tanner graph \mathcal{G} , as shown in Fig. 3. The proposed improved decoding algorithm is based on the joint Tanner graph \mathcal{G} , and the major alterations are the VNs of \mathcal{G}_1 update and CNs of \mathcal{G}_2 update.

- VNs of \mathcal{G}_1 update. Since some of the VNs of \mathcal{G}_1 are connected by both the CNs of \mathcal{G}_1 and CNs of \mathcal{G}_2 , the update LLRs of these VNs should taken into consideration both CNs set.
- CNs of \mathcal{G}_2 update. For the CNs of \mathcal{G}_2 , it connects both its own VNs and some VNs of \mathcal{G}_1 , thus, the update LLRs of the CNs of \mathcal{G}_2 are determined by the two VNs sets.

In the algorithm, there are some definitions of MPA. The VN and CN nodes of \mathcal{G}_1 are defined by α_1 and β_1 ; similarly, the VN and CN nodes of \mathcal{G}_2 are set to be α_2 and β_2 . Let $N(\alpha_1)$, $N(\alpha_2)$, $N(\beta_1)$ and $N(\beta_2)$ be the sets that are respectively connected with VN node α_1 , VN node α_2 , CN node β_1 and CN node β_2 . Note that, there is no edge between α_2 and β_1 , because of the SPC structure.

In summary, the receiver should detect the number of arrival users, separate the superimposed signals, and recover the transmit signals of all the users. The first step is to do SoF-AUD, followed by the MUD and improved MPA steps. It is noted that, the initial soft LLR information of MPA is decided by the MUD step. In other words, there is a close relationship between MUD and MPA, sometimes, the MUD and MPA can be jointly considered. The entire detection algorithms are summarized in Algorithm 1.

V. THEORETICAL ANALYSIS

In this section, we deduce theoretical AUER and Shannon limits for the proposed system. The AUER is utilized to measure the performance of AUD. Moreover, Shannon limits of the proposed system are derived, instead of the theoretical BER.

A. AUER

For a random access system, define active user error rate (AUER) by

$$P_{e,AU} = \Pr[\tau \neq \hat{\tau}], \quad (19)$$

Algorithm 1: Detection algorithms of the GFMA receiver.**Input:** \mathbf{y}_m for $1 \leq m \leq M$, N_{max} .**Output:** $\mathbf{u}^{(j)}$ for $1 \leq j \leq J$.**Step 1: SoF-AUD:** Obtain $\hat{\tau}$ and $\hat{\mu}$ by (13);**Step 2: MUD****if without encoder \mathbf{G}_1 then** Obtain $\hat{\omega}_{m,l,1}$ by (14) when $\mathcal{M}_1 = 1$; and $\hat{\omega}_{m,l}$ by (15) when $\mathcal{M}_1 > 1$. **Output:** Hard-decoding for $\mathbf{c}_m^{(j)}$, then obtain $\mathbf{u}^{(j)}$;**end**Obtain initial LLRs of α_1 and α_2 by (18);**Step 3: Improved MPA**Set $\text{LLR}_{\alpha_1 \rightarrow \beta_1 (\text{or } \beta_2)} = \text{LLR}_{\alpha_1}$; $\text{LLR}_{\alpha_2 \rightarrow \beta_2} = \text{LLR}_{\alpha_2}$; Set $N_{sim} = 0$;**while** $N_{sim} < N_{max}$ **do** **Update CNs:**

$$\text{LLR}_{\beta_1 \rightarrow \alpha_1} = 2 \tanh^{-1} \left[\prod_{\alpha'_1 \in N(\beta_1) - \alpha_1} \tanh \left(\frac{1}{2} \text{LLR}_{\alpha'_1 \rightarrow \beta_1} \right) \right];$$

$$\text{LLR}_{\beta_2 \rightarrow \alpha_1} = 2 \tanh^{-1} \left[\prod_{\alpha'_1 \in N(\beta_2) - \alpha_1} \tanh \left(\frac{1}{2} \text{LLR}_{\alpha'_1 \rightarrow \beta_2} \right) \right];$$

$$\text{LLR}_{\beta_2 \rightarrow \alpha_2} = 2 \tanh^{-1} \left[\prod_{\alpha'_2 \in N(\beta_2) - \alpha_2} \tanh \left(\frac{1}{2} \text{LLR}_{\alpha'_2 \rightarrow \beta_2} \right) \right];$$

Update VNs:

$$\text{LLR}_{\alpha_1 \rightarrow \beta_1} = \text{LLR}_{\alpha_1} + \sum_{\beta'_1 \in N(\alpha_1) - \beta_1} L_{\beta'_1 \rightarrow \alpha_1};$$

$$\text{LLR}_{\alpha_1 \rightarrow \beta_2} = \text{LLR}_{\alpha_1} + \sum_{\beta'_2 \in N(\alpha_1) - \beta_2} L_{\beta'_2 \rightarrow \alpha_1}; \text{LLR}_{\alpha_2 \rightarrow \beta_2} = \text{LLR}_{\alpha_2};$$

Update LLR:

$$\text{LLR}_{\alpha_1} = \text{LLR}_{\alpha_1} + \sum_{\beta \in N(\alpha_1)} \text{LLR}_{\beta_1 \rightarrow \alpha_1} + \sum_{\beta \in N(\alpha_2)} \text{LLR}_{\beta_2 \rightarrow \alpha_1};$$

 Set $v_{n_1}^{(j)} = 1$ if $\text{LLR}_{\alpha_1} > 0$, otherwise, $v_{n_1}^{(j)} = 0$; **if** $N_{sim} \geq N_{max}$ **||** $\mathbf{v}^{(j)} \cdot \mathbf{H}_1 = \mathbf{0}$ **then** **Output:** $\mathbf{v}^{(j)} \rightarrow \mathbf{u}^{(j)}$; **end****end**

where τ and $\hat{\tau}$ respectively stand for the factual and detected numbers of arrival users.

In fact, the AUER of our proposed system is mainly determined by two aspects, one is the detection error caused by noise, and the other is the error floor caused by insufficient information. Our proposed detection algorithm is based on the sum power γ_{sum} , which is affected by several

parameters that are the sum-pattern set of τ users Ω^τ , the number of rows M , and multi-dimensional order \mathcal{M}_1 . Thus, the AUER can be derived as

$$P_{e,AU} = \sum_{\tau=1}^T P_\tau \sum_{1 \leq \tau' \leq T, \tau' \neq \tau} \int_{D_{\tau'}} p(\gamma_{sum}|\tau) d\gamma_{sum}, \quad (20)$$

where P_τ is *a priori* probability of the number of arrival users, $p(\gamma_{sum}|\tau)$ is the conditional probability density function (PDF) of γ_{sum} given by the number of arrival users τ for $\tau = 1, 2, \dots, T$, and $D_{\tau'}$ is the decision region of τ' .

Thereafter, the key issue is to derive the PDF of $p(\gamma_{sum}|\tau)$. Regarding as the sum-form of γ_{sum} as shown in (11), $p(\gamma_{sum}|\tau)$ is determined by a sequence of noncentral Chi-Square random variables, defined by

$$p(\gamma_{sum}|\tau) = \sum_{s_a \in \mathcal{S}} P(s_a) \cdot f_{\chi^2}(s_a), \quad (21)$$

where $f_{\chi^2}(s_a)$ is the PDF of a noncentral Chi-Square random variable γ_s , where $\gamma_s = M \cdot \gamma_{sum}$, given by

$$f_{\chi^2}(s_a) = \frac{1}{N_0} \left(\frac{\gamma_s}{s_a^2} \right)^{\frac{N_f-2}{4}} e^{-\frac{s_a^2 + \gamma_s}{N_0}} \cdot I_{\frac{N_f}{2}-1} \left(\frac{s_a}{N_0/2} \sqrt{\gamma_s} \right), \quad (22)$$

where $I_\alpha(r) = \sum_{t=0}^{\infty} \frac{(r/2)^{\alpha+2t}}{t! \Gamma(\alpha+t+1)}$ and $\Gamma(r) = \int_0^{\infty} t^{r-1} e^{-t} dt$. γ_s is consisted of N_f independent Gaussian variables with comment variance $N_0/2$ and different means denoted by s_{n_f} . Since $r_{m,l,i}$ is a complex number, the degree of freedom is $N_f = 2L \cdot M$.

In (21), s_a is defined by $s_a = \sqrt{\sum_{n_f=1}^{N_f} s_{n_f}^2}$, and all the values of s_a are from the set \mathcal{S} , i.e., $s_a \in \mathcal{S}$. The probability of s_a in set \mathcal{S} is defined by $P(s_a)$, thus, $\sum_{s_a \in \mathcal{S}} P(s_a) = 1$.

Then, analyzing $p(\gamma_{sum}|\tau)$ is equivalent to finding the values of s_a and $P(s_a)$, which can be calculated by the sum-pattern set Ω^τ . Let us review the expression of Ω^τ , which is defined by $\Omega^\tau = \{\Omega_1^\tau, \Omega_2^\tau, \dots, \Omega_l^\tau, \dots, \Omega_L^\tau\}$ with $\Omega_l^\tau = \omega_l^{(\tau,1)} \cup \dots \cup \omega_l^{(\tau,\mu)} \cup \dots \cup \omega_l^{(\tau,C_T^\tau)}$.

Divide the complex number set $\omega_l^{(\tau,\mu)}$ into two sets, the real number set $\omega_{l,re}^{(\tau,\mu)}$ and the image number set $\omega_{l,im}^{(\tau,\mu)}$. Evidently, s_{n_f} belongs to the set $\omega_{l,re}^{(\tau,\mu)} \cup \omega_{l,im}^{(\tau,\mu)}$, i.e., $s_{n_f} \in \omega_{l,re}^{(\tau,\mu)} \cup \omega_{l,im}^{(\tau,\mu)}$, varying with the varied τ and μ , so to the variable s_a . For example, if $\tau = 1$, it is known that $s_a = \sqrt{\sum_{n_f=1}^{N_f} s_{n_f}^2} = \sqrt{M \sum_{l=1}^L a_l^2} = \sqrt{ML \cdot P_{avg}}$.

In general, s_a and $P(s_a)$ are both affected by the ratio of the number of elements in $\omega_l^{(\tau,\mu)}$ to the total number of elements in Ω_l^τ . Therefore, different UDAS sets result in different s_a and $P(s_a)$, thus obtaining different $p(\gamma_{sum}|\tau)$.

To explain the analysis processing of $p(\gamma_{sum}|\tau)$, we present an example. If we exploit the UDAS set given by Example 2, as shown in (5), we have

$$\begin{aligned}
 p(\gamma_{sum}|\tau=1) &= f_{\chi^2}(\sqrt{ML \cdot P_{avg}}), \\
 p(\gamma_{sum}|\tau=2) &= \frac{2}{3}f_{\chi^2}(\sqrt{20M}) + \frac{1}{3} \sum_{k=0}^{LM} \left[p(k, \frac{N_f}{2}) f_{\chi^2}(\sqrt{9k + (LM - k)}) \right], \\
 p(\gamma_{sum}|\tau=3) &= \sum_{k=0}^{LM} \left[p(k, \frac{N_f}{2}) f_{\chi^2}(\sqrt{9k + (LM - k) + 10M}) \right], \\
 p(\gamma_{sum}|\tau=4) &= \sum_{k=0}^{N_f} \left[p(k, N_f) f_{\chi^2}(\sqrt{9k + (N_f - k)}) \right],
 \end{aligned} \tag{23}$$

where $p(k, N_f) = C_{N_f}^k \cdot (\frac{1}{2})^{N_f}$ is the probability of a binomial distribution that selects k from N_f . The case of $\tau = 1$ is accord with the aforementioned discussion.

When $\tau = 2$, it is known that μ belongs to $[1, C_T^T]$, and $\Psi^{(2,\mu)}$ is given in Example 4. Specifically, they can be divided into the following two cases.

- 1) Case 1: $\mu = 1, 3, 4$ or 6 . At the moment, the sum-patterns of the two users include both real and image numbers. Take $\mu = 1$ as an example, we can get $\omega_1^{(2,1)} = \{1 + 2i, 1 - 2i, -1 + 2i, -1 - 2i\}$, $\omega_2^{(2,1)} = \{1 + 1i, 1 - 1i, -1 + 1i, -1 - 1i\}$, $\omega_3^{(2,1)} = \{2 + 1i, 2 - 1i, -2 + 1i, -2 - 1i\}$ and $\omega_4^{(2,1)} = \{2 + 2i, 2 - 2i, -2 + 2i, -2 - 2i\}$. Therefore, the power of the sum-pattern always keeps as a constant. For example, the power values of the four symbols ($l = 1, 2, 3, 4$) are respectively 5, 2, 5, and 8. Then, s_a is calculated as $s_a = \sqrt{M \cdot (5 + 2 + 5 + 8)} = \sqrt{20M}$. The cases of $\mu = 3, 4$ and 6 are exactly the same as the case of $\mu = 1$.
- 2) Case 2: $\mu = 2$ or 5 . At the moment, the sum-patterns of the two users are all real numbers (or image numbers). Obviously, we can get $\Omega^{(2,2)} = \{+3, +1, -1, -3\}$ and $\Omega^{(2,5)} = \{+3i, +1i, -1i, -3i\}$. Now, we take $\mu = 2$ as an example to analyze. Since the sum-patterns of $\mu = 2$ are all real numbers, then LM Gaussian variables of the totally N_f degrees are distributed as $\mathcal{N}(0, N_0/2)$, and the others LM variables are in probability distributed as $\mathcal{N}(\pm 3, N_0/2)$ or $\mathcal{N}(\pm 1, N_0/2)$. Assume k Gaussian variables are distributed as $\mathcal{N}(\pm 3, N_0/2)$, and the rest $LM - k$ Gaussian variables are with distribution as $\mathcal{N}(\pm 1, N_0/2)$, where $0 \leq k \leq LM$ with probability $C_{LM}^k (\frac{1}{2})^{LM}$. Thereafter, s_a is calculated as $s_a = \sqrt{k \cdot 9 + (LM - k) \cdot 1 + LM \cdot 0} = \sqrt{9k + (LM - k)}$.

Consider the probabilities of cases 1 and 2 are respectively $2/3$ and $1/3$, we can derive the final PDF of $p(\gamma_{sum}|\tau = 2)$ as shown in (23). When $\tau > 2$, the analysis processing is the same as

the case of $\tau = 2$, which will not be repeated described.

It is remarkable that the detection error probability of the UDAS set $\Psi^{(\tau,\mu)}$ can be ignored compared to the AUER.

B. Shannon limit of the proposed system in an adder multiple-access channel

The computation of Shannon limit involves two steps, calculating the channel capacity and solving the lowest E_b/N_0 , where E_b represents the bit energy of each user.

Assume R_j is the data rate of the j th user, and P_j is the transmit power of the j th user. Then, the capacity of a MAC is given by [30]

$$\sum_{j=1}^J R_j \leq \log_2 \left(1 + \frac{P_1 + P_2 + \dots + P_J}{N_0} \right). \quad (24)$$

If all the users hold the same data rate, i.e., $R_1 = R_2 = \dots = R_J = R_c$, the last inequality dominates the others. It is remarkable that the Shannon limit of the proposed system is deduced based on this assumption. Besides, due to the cyclic UDAS set, it is able to known $P_1 = P_2 = \dots = P_J = P_{avg}$. Thus, the relationship between P_{avg} and E_b is $P_{avg} = \frac{R_c \cdot \log_2(\mathcal{M}) \cdot E_b}{T_s}$, where T_s is the symbol duration.

In the following discussion, assume the number of arrival users $J = \tau$ and the selected UDAS set $\Psi^{(\tau,\mu)}$, i.e., $\Psi^{(\tau,\mu)} = \{\mathbf{e}_{t_1}, \mathbf{e}_{t_2}, \dots, \mathbf{e}_{t_\nu}, \dots, \mathbf{e}_{t_\tau}\}$ with $t_\nu \in \{1, 2, \dots, T\}$, have been available. Denote $\{\vec{\chi}_{l,m_s}^{(j)}, 1 \leq m_s \leq \mathcal{M}\}$ by the transmit constellation set of the l th symbol of the j th user, including \mathcal{M} possible \mathcal{M}_1 -dimension transmit signals, expressed as $\vec{\chi}_{l,m_s}^{(j)} = \{\chi_{l,m_s,1}^{(j)}, \chi_{l,m_s,2}^{(j)}, \dots, \chi_{l,m_s,i}^{(j)}, \dots, \chi_{l,m_s,\mathcal{M}_1}^{(j)}\}$, where $\chi_{l,m_s,i}^{(j)} \in \{-e_{t_j,l}, 0, e_{t_j,l}\}$, with the assumption that the j th user selects the UDAS \mathbf{e}_{t_j} in $\Psi^{(\tau,\mu)}$. Define $\Pr(\vec{\chi}_{l,m_s}^{(j)})$ by the *prior* probability of $\vec{\chi}_{l,m_s}^{(j)}$.

When J users' signals, i.e., $\vec{\chi}_{l,m_s}^{(j)}$, are transmitted through a MAC, the PDF of the l th symbol of the received signal $\mathbf{y}_{m,l}$ is

$$p(\mathbf{y}_{m,l} | \vec{\chi}_{l,m_s}^{(1)}, \vec{\chi}_{l,m_s}^{(2)}, \dots, \vec{\chi}_{l,m_s}^{(J)}) = \prod_{m_s=1}^{\mathcal{M}} \dots \prod_{m_s=1}^{\mathcal{M}} \frac{1}{\sqrt{\pi N_0}} \exp \left\{ -\frac{(\mathbf{y}_{m,l} - \sum_{j=1}^J \vec{\chi}_{l,m_s}^{(j)})^2}{N_0} \right\}. \quad (25)$$

Note that the row index has no effect to the PDF.

TABLE I

SHANNON LIMITS OF THE PROPOSED SYSTEM IN AN ADDER MULTIPLE-ACCESS CHANNEL, WHERE $\mathcal{M}_1 = 1$, $J = 2$, $L = 4$, AND $\Psi^{(2,1)} = \{(1, 1i, 2, 2i), (2i, 1, 1i, 2)\}$. THE UNIT OF $(\frac{E_b}{N_0})_{\min}$ IS DB.

R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$
0.001	-1.5858	0.151	-0.9594	0.435	0.4667	0.646	2.0022	0.794	3.6701	0.895	5.4437
0.006	-1.5682	0.175	-0.8524	0.459	0.6154	0.663	2.1577	0.804	3.8150	0.901	5.5703
0.013	-1.5392	0.226	-0.6215	0.483	0.7657	0.694	2.4673	0.814	3.9586	0.921	6.0639
0.022	-1.4992	0.250	-0.4985	0.506	0.9176	0.709	2.6212	0.823	4.1007	0.934	6.4212
0.054	-1.3716	0.305	-0.2394	0.550	1.2248	0.736	2.9263	0.840	4.3807	0.951	6.9939
0.066	-1.3185	0.330	-0.1042	0.571	1.3797	0.749	3.0774	0.856	4.6550	0.962	7.4329
0.106	-1.1538	0.384	0.1759	0.610	1.6907	0.773	3.3762	0.877	5.0556	0.981	8.4635
0.128	-1.0600	0.416	0.3566	0.628	1.8465	0.784	3.5238	0.883	5.1864	0.991	9.3165

Consider the transmit mode of each user, the channel capacity of the l th symbol in the adder MAC can be calculated as

$$C_l = \max_{\mathcal{M}\text{-fold}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(\mathbf{y}_{m,l} | \vec{\chi}_{l,m_s}^{(1)}, \vec{\chi}_{l,m_s}^{(2)}, \dots, \vec{\chi}_{l,m_s}^{(J)}) \cdot \Pr(\vec{\chi}_{l,m_s}^{(1)}) \cdots \Pr(\vec{\chi}_{l,m_s}^{(J)}) \cdot \log_2 \left[\frac{p(\mathbf{y}_{m,l} | \vec{\chi}_{l,m_s}^{(1)}, \vec{\chi}_{l,m_s}^{(2)}, \dots, \vec{\chi}_{l,m_s}^{(J)})}{p(\mathbf{y}_{m,l})} \right] d\mathbf{y}_{m,l}. \quad (26)$$

where \mathcal{Y} is the the received signal's variable, $\mathcal{X}_1, \mathcal{X}_2, \dots$, and \mathcal{X}_J are the transmit signals' variables of J users.

Since each user exploits an L -length UDAS, we define ergodic capacity as

$$\bar{C} = E[C_l] = \frac{1}{L} \sum_{l=1}^L C_l. \quad (27)$$

Regarding as (24), it is known that

$$R_c \cdot \log_2 \mathcal{M} \cdot J \leq \bar{C} = f(E_b/N_0). \quad (28)$$

When $R_c \cdot \log_2 \mathcal{M} \cdot J = \bar{C}$, the utilization of communication resources is maximized. The Shannon limit is defined as the minimum E_b/N_0 that can realize reliable transmission with R_c , which is denoted by $(E_b/N_0)_{\min}$. Thus,

$$(E_b/N_0)_{\min} = f^{-1}(R_c \cdot \log_2 \mathcal{M} \cdot J). \quad (29)$$

However, the integral in (26) is extremely complex, especially the inverse function. Since it is one-to-one mapping between the code rate R_c and $(E_b/N_0)_{\min}$, the corresponding code rate

TABLE II

SHANNON LIMITS OF THE PROPOSED SYSTEM IN AN ADDER MULTIPLE-ACCESS CHANNEL, WHERE $\mathcal{M}_1 = 1$, $J = 3$, $L = 4$,
 $\Psi^{(3,1)} = \{(1, 1i, 2, 2i), (2i, 1, 1i, 2), (2, 2i, 1, 1i)\}$. THE UNIT OF $(\frac{E_b}{N_0})_{\min}$ IS DB.

R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$
0.001	-1.5840	0.214	-0.3740	0.499	1.6466	0.690	3.4298	0.813	4.8805	0.907	6.4123
0.006	-1.5609	0.237	-0.2279	0.518	1.8023	0.703	3.5688	0.822	5.0045	0.912	6.5236
0.013	-1.5229	0.261	-0.0785	0.536	1.9571	0.716	3.7063	0.830	5.1274	0.922	6.7438
0.022	-1.4706	0.284	0.0736	0.554	2.1106	0.728	3.8423	0.847	5.3699	0.931	6.9608
0.049	-1.3268	0.330	0.3839	0.587	2.4139	0.751	4.1100	0.862	5.6081	0.953	7.5932
0.065	-1.2371	0.353	0.5411	0.604	2.5635	0.763	4.2417	0.869	5.7258	0.956	7.6959
0.103	-1.0277	0.397	0.8575	0.634	2.8583	0.784	4.5011	0.883	5.9582	0.971	8.2967
0.167	-0.6532	0.439	1.1745	0.663	3.1471	0.803	4.7552	0.895	6.1870	0.990	9.5969

TABLE III

SHANNON LIMITS OF THE PROPOSED SYSTEM IN AN ADDER MULTIPLE-ACCESS CHANNEL, WHERE $\mathcal{M}_1 = 1$, $J = 2$, $L = 6$,
 AND $\Psi^{(2,1)} = \{(1, 1i, 2, 2i, 4, 4i), (4i, 1, 1i, 2, 2i, 4)\}$. THE UNIT OF $(\frac{E_b}{N_0})_{\min}$ IS DB.

R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$	R_c	$(\frac{E_b}{N_0})_{\min}$
0.004	-1.5693	0.275	0.0848	0.574	2.9037	0.726	5.1081	0.914	9.1845	0.833	7.0888
0.016	-1.5033	0.312	0.3579	0.595	3.1718	0.739	5.3254	0.918	9.3288	0.841	7.2666
0.035	-1.3973	0.348	0.6377	0.615	3.4343	0.752	5.5379	0.922	9.4709	0.848	7.4413
0.090	-1.0846	0.415	1.2092	0.651	3.9415	0.775	5.9492	0.941	10.1529	0.863	7.7819
0.124	-0.8872	0.446	1.4972	0.668	4.1862	0.786	6.1485	0.950	10.5407	0.876	8.1111
0.198	-0.4309	0.502	2.0699	0.698	4.6582	0.806	6.5354	0.971	11.6227	0.894	8.5855
0.236	-0.1790	0.529	2.3521	0.713	4.8858	0.815	6.7233	0.981	12.2853	0.904	8.8895

R_c can be calculated for a given $(E_b/N_0)_{\min}$. Then, we can obtain a sequence of data points $(R_c, (E_b/N_0)_{\min})$. In practice, the range of the integral and the accuracy of interpolation are depended on the required accuracy of the Shannon limit.

Table I, Table II and Table III show the Shannon limits of the proposed system in an adder multiple-access channel. When R_c and L are given, it is found that the required $(E_b/N_0)_{\min}$ of the $J = 3$ case is larger than the $J = 2$ case. Moreover, for a given number of arrival users J and a given R_c , a larger L indicates a larger $(E_b/N_0)_{\min}$, because of the reduced minimum distance of the constellation at the receiver.

VI. SIMULATION RESULTS

In this section, we simulate and compare the performances of the proposed system. The UDAS sets are generated based on the cyclic matrix mode, with generators $\mathbf{a} = (1, 1i, 2, 2i)$ of $L = 4$,

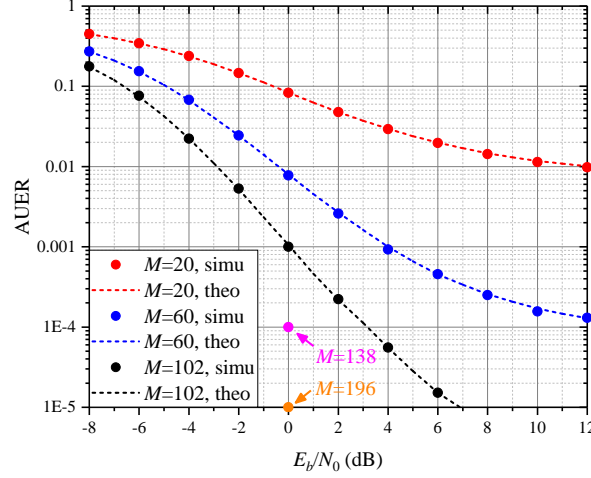


Fig. 4. AUER of the proposed system with various parameters, where $\mathcal{M}_1 = 1$, $L = 4$, and $M = 20, 60, 102, 138$ and 196 .

and $\mathbf{a} = (1, 1i, 2, 2i, 4, 4i)$ of $L = 6$. Moreover, the first encoder exploits a (3,6)-regular QC-LDPC (1016,508) with rate 0.5, and the second encoder is a SPC with rate $R_2 = 1 - \frac{1}{L \cdot \log_2 \mathcal{M}}$. Thus, the total data rate of each user is equal to $R_c = 0.5 \times (1 - \frac{1}{L \cdot \log_2 \mathcal{M}})$. For example, when $L = 4$ and $\mathcal{M} = 2$, it is found that $R_c = 0.375$.

First of all, we focus on the AUER performance, with the assumption that P_τ is uniform distribution in $[1, T]$. To observe the AUER performance, we set $R_c = 1$ and omit the two encoders. The AUER performance of the proposed system is shown in Fig. 4.

It can be seen from Fig. 4 that the AUER decreases with the increased E_b/N_0 . When M is small, e.g., $M = 20$, there exists a significantly error floor, because of the lack of statistical information. With the increase of M , the AUER improves significantly, revealing that sufficient statistical information may reduce both the influence of noise and error floor. Moreover, the simulated results are perfect accord with our deduced theoretical results.

When we set $E_b/N_0 = 0$ dB and $\text{AUER} \approx 10^{-3}, 10^{-4}$ and 10^{-5} , it is found that the minimum required numbers of M are respectively 102, 138, and 196, whose corresponding transmit blocks are with length 408, 552 and 784. Evidently, these block lengths satisfy the requirement of short packet communications. Therefore, we can maximum simultaneously support 4 users with a 10^{-5} detected error rate, at this moment, the required E_b/N_0 is only 0 dB. This result is appealing for a random access network.

The BER performances of the proposed system are shown in Fig. 5, where the block length of the transmit block is equal to or larger than 1016 that is the codeword length of the first encoder.

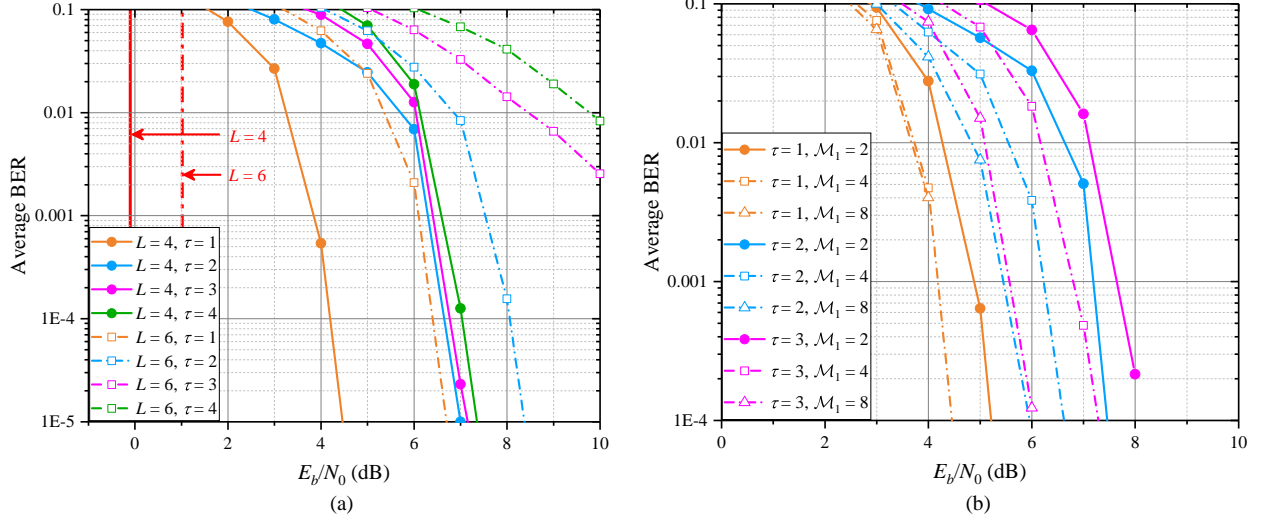


Fig. 5. Average BER performances of the proposed system, where the maximum iteration number is 10. (a) $M_1 = 1$ case, with $L = 4, 6$, and $\tau = 1, 2, 3, 4$. (b) $M_1 > 1$ case, with $L = 4$, and $\tau = 1, 2, 3$.

Therefore, we assume that the number of arrival users and UDAS set have been perfectly detected.

Fig. 5(a) shows the BER performance of the proposed scheme with parameter $M_1 = 1$. When $L = 4$, it is found that BER of the case $\tau = 1$ provides the best performance, following by the cases of $\tau = 2$, $\tau = 3$ and $\tau = 4$. However, the BER differences among $\tau = 2, 3, 4$ are small. When BER is 10^{-4} and $L = 4$, the BER of $\tau = 3$ is about 6.8 dB away from the Shannon limit, and 6.7 dB away from the Shannon limit for the case of $\tau = 2$. Therefore, there exists an improvement space to further reduce the BER. For a given τ , e.g., $\tau = 2$, the BER of the case of $L = 4$ is much better than the case of $L = 6$, since the received superimposed signal of the $L = 4$ case has a relative large minimum Euclidean distance. For example, when $\tau = 2$ and BER is 10^{-4} , the case of $L = 4$ is about 1.5 dB better than that of the case of $L = 6$. This result is accord with our deduced Shannon limits.

The BER performance of the proposed scheme with various dimensions $M_1 = 2, 4, 8$ is shown in Fig. 5(b). It can be seen that the BER performance is improved with the increased M_1 , due to the statistic feature and larger Euclidean distance. For example, when $\tau = 2$ and BER is 10^{-3} , the case of $M_1 = 4$ provides about 0.7 dB gain than the case of $M_1 = 2$.

Fig. 6 shows the BER comparisons among different systems. We compare our proposed system with the classical CDMA system that exploits Walsh sequences whose SF is 4, and the SCMA system that is based on the codebook given by [19]. Therefore, the SEs of the CDMA and SCMA

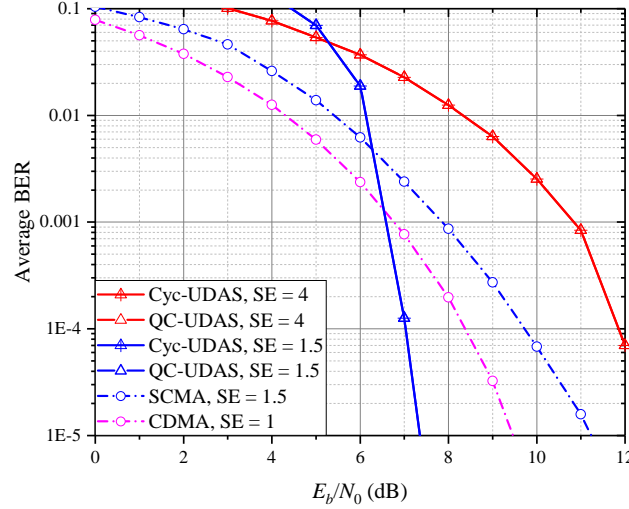


Fig. 6. BER comparisons among different systems. The CDMA system utilizes Walsh sequences whose SF is 4, and SCMA system is operated based on the codebook given by [19]. Both the schemes one and two of our proposed systems have the same parameters, where $L = 4$, $\mathcal{M}_1 = 1$, and $\tau = 4$. The scheme one is without any encoder; and scheme two is with two encoders. Both cyclic and quasi-cyclic structures are taken into consideration.

systems are respectively 1 bit/resource and 1.5 bits/resource. Our proposed system considers two schemes for comparisons. The scheme one is corresponding to the parameters of $L = 4$, $\mathcal{M}_1 = 1$, $\tau = 4$ without any encoders, indicating that the SE of case one is equal to 4 bits/resource. On the other hand, the scheme two has the same parameters as that of the scheme one, and the only difference is caused by the inserted encoders, resulting in the SE of scheme two becomes as 1.5 bits/resources.

It is found that, the BER of our proposed scheme one is worse than those of CDMA and SCMA systems. However, our proposed scheme one provides much higher SE. Based on the inserted channel codes, our proposed scheme two can provide much better BER performances than those of CDMA and SCMA, when E_b/N_0 is larger than 6.6 dB. At this time, our proposed scheme two has the same SE as that of SCMA system, verifying the validity of our proposed system. Moreover, both the cyclic and quasi-cyclic UDAS modes provide the same BER performances, since they have the same UDAS set.

VII. CONCLUSION

This paper introduces UDAS for grant-free multiple-access systems. First of all, we present an UDAS-based MD-BICM transmitter for each user, which is a combination of two channel encoders, one interleaver and multi-dimensional modulation. The first encoder is used to improve

the reliability of the system, and the second encoder is used for assisting AUD. Following, some definitions of UDAS are introduced in details. Refer to the constructions of QC-LDPC, two kinds of UDAS sets are presented, which are cyclic and quasi-cyclic matrix modes. The cyclic/quasi-cyclic structure can help the receiver realize low-complexity AUD. Thirdly, we present a SoF-AUD, and a joint MUD and improved MPA for the proposed system, where the Tanner graph of the decoder is the combination of two encoders \mathcal{G}_1 and \mathcal{G}_2 . Finally, both the theoretical AUER and Shannon limits are deduced in details. We simulate and compare our proposed system with various parameters and systems. When $E_b/N_0 = 0$ dB and the length of transmit block is larger than a given value, e.g., 784, the AUER of our proposed system can be an extremely low value 10^{-5} , and the system can maximum detect four simultaneously arrival users. Essentially, we can design the parameters of a UDAS-based transmitter, to satisfy the requirement of a GFMA system. Our proposed system can provide high spectrum efficiency, which can compare with a designed NOMA codebook. For example, when $E_b/N_0 = 7$ dB and $SE = 1.5$, the BER of our proposed system is about 10^{-4} , which is much better than that of the SCMA system.

Actually, the proposed UDAS-based MD-BICM transmitter can be extended to a general random access scenario, e.g., uncoordinated multiple-access. Moreover, there are many approaches to construct various UDAS sets. These contents are left as the future works.

REFERENCES

- [1] L. Dai, B. Wang, Y. Yuan, S. Han, C. I. and Z. Wang, "Nonorthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Communication Magazine*, vol. 53, no. 9, pp. 74-81, Sept. 2015.
- [2] 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 Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2181-2195, July 2017.
- [3] J. Dai, K. Niu, Z. Si, C. Dong and J. Lin, "Polar-Coded Non-Orthogonal Multiple Access," *IEEE Transactions on Signal Processing*, vol. 66, no. 5, pp. 1374-1389, March 2018.
- [4] Y. Wu, X. Gao, S. Zhou, W. Yang, Y. Polyanskiy and G. Caire, "Massive Access for Future Wireless Communication Systems," *IEEE Wireless Communications*, vol. 27, no. 4, pp. 148-156, August 2020.
- [5] J. Belfiore, G. Rekaya, and E. Viterbo, "The golden code: A 2×2 fullrate space-time code with nonvanishing determinants," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1432-1436, Apr. 2005.
- [6] H. Zhang, R. Li, J. Wang, Y. Chen and Z. Zhang, "Reed-Muller Sequences for 5G Grant-Free Massive Access," *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017, pp. 1-7.
- [7] O. Mashayekhi and F. Marvasti, "Uniquely Decodable Codes with Fast Decoder for Overloaded Synchronous CDMA Systems," *IEEE Transactions on Communications*, vol. 60, no. 11, pp. 3145-3149, Nov. 2012.
- [8] A. Singh and P. Singh, "Uniquely decodable codes for overloaded synchronous CDMA with two sets of orthogonal signatures," *2014 Annual IEEE India Conference (INDICON)*, 2014, pp. 1-4.

- [9] M. Kulhandjian, H. Kulhandjian, C. D'Amours, H. Yanikomeroglu and G. Khachatrian, "Fast Decoder for Overloaded Uniquely Decodable Synchronous Optical CDMA," *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, 2019, pp. 1-7.
- [10] M. Kulhandjian and D. A. Pados, "Uniquely decodable code-division via augmented Sylvester-Hadamard matrices," *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, 2012, pp. 359-363.
- [11] M. Li and Q. Liu, "Fast Code Design for Overloaded Code-Division Multiplexing Systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 1, pp. 447-452, Jan. 2016.
- [12] M. Kulhandjian, C. D'Amours and H. Kulhandjian, "Uniquely Decodable Ternary Codes for Synchronous CDMA Systems," *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2018, pp. 1-6.
- [13] D. V. Sarwate and M. B. Pursley, "Cross correlation properties of pseudo random and related sequences," *Proceedings of the IEEE*, vol. 68, no. 5, pp. 593-619, May 1980.
- [14] B. M. Popovic, "Generalized chirp-like polyphase sequences with optimum correlation properties," *IEEE Transactions on Information Theory*, vol. 38, no. 4, pp. 1406-1409, July 1992.
- [15] M. Hua et al., "Analysis of the Frequency Offset Effect on Random Access Signals," *IEEE Transactions on Communications*, vol. 61, no. 11, pp. 4728-4740, November 2013.
- [16] Y. Tsai, H. Huang, Y. Chen and K. Yang, "Simultaneous Multiple Carrier Frequency Offsets Estimation for Coordinated Multi-Point Transmission in OFDM Systems," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4558-4568, September 2013.
- [17] R. Abbas, M. Shirvanimoghaddam, Y. Li and B. Vucetic, "A Novel Analytical Framework for Massive Grant-Free NOMA," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2436-2449, March 2019.
- [18] A. Bayesteh, E. Yi, H. Nikopour and H. Baligh, "Blind detection of SCMA for uplink grant-free multiple-access," *2014 11th International Symposium on Wireless Communications Systems (ISWCS)*, 2014, pp. 853-857.
- [19] Y. Zhou, Q. Yu, W. Meng and C. Li, "SCMA codebook design based on constellation rotation," *2017 IEEE International Conference on Communications (ICC)*, 2017, pp. 1-6.
- [20] H. Huang, H. Zhao, F. Wang and J. Li, "Uplink Grant-free Multi-codebook SCMA Based on High-overload Codebook Grouping," *2018 IEEE/CIC International Conference on Communications in China (ICCC Workshops)*, 2018, pp. 112-116.
- [21] X. Zhang, G. Han, D. Zhang and L. Yang, "A Lattice-Based SCMA Codebook Design for IoMT Communications," *2019 IEEE/CIC International Conference on Communications Workshops in China (ICCC Workshops)*, 2019, pp. 169-173.
- [22] L. Yu, P. Fan, X. Lei and P. T. Mathiopoulos, "BER Analysis of SCMA Systems With Codebooks Based on Star-QAM Signaling Constellations," *IEEE Communications Letters*, vol. 21, no. 9, pp. 1925-1928, Sept. 2017.
- [23] L. Yu, P. Fan, D. Cai and Z. Ma, "Design and Analysis of SCMA Codebook Based on Star-QAM Signaling Constellations," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10543-10553, Nov. 2018.
- [24] B. Wang, L. Dai, Y. Zhang, T. Mir and J. Li, "Dynamic Compressive Sensing-Based Multi-User Detection for Uplink Grant-Free NOMA," *IEEE Communications Letters*, vol. 20, no. 11, pp. 2320-2323, Nov. 2016.
- [25] Y. Du et al., "Efficient Multi-User Detection for Uplink Grant-Free NOMA: Prior-Information Aided Adaptive Compressive Sensing Perspective," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2812-2828, Dec. 2017.
- [26] L. Liu, E. G. Larsson, W. Yu, P. Popovski, C. Stefanovic and E. de Carvalho, "Sparse Signal Processing for Grant-Free Massive Connectivity: A Future Paradigm for Random Access Protocols in the Internet of Things," *IEEE Signal Processing Magazine*, vol. 35, no. 5, pp. 88-99, Sept. 2018.
- [27] Q. Yu, W. Meng and S. Lin, "Packet Loss Recovery Scheme with Uniquely-Decodable Codes for Streaming Multimedia over P2P Networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 142-154, September 2013.

- [28] Q. Yu, H. Chen and W. Meng, "A Unified Multiuser Coding Framework for Multiple-Access Technologies," *IEEE Systems Journal*, vol. 13, no. 4, pp. 3781-3792, Dec. 2019.
- [29] J. Li, S. Lin, K. Abdel-Ghaffar, W. E. Ryan, and D. J. Costello, *LDPC Code Designs, Constructions, and Unification*, Cambridge, UK: Cambridge Univ Press, 2017.
- [30] Thomas M. Cover, and Joy A. Thomas, *Elements of Information Theory*, Tsinghua University Press, 2002.