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

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

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) without additional redundancy. We give the definition of a $T$-size UDAS set, and construct two kinds of UDAS sets based on cyclic and quasi-cyclic modes. Then, we propose an UDAS-based multi-dimensional bit interleaving coded modulation (MD-BICM) transceiver, with corresponding detection algorithms, i.e., a statistic of UDAS feature based active user detection algorithm (SoF-AUD) and a multiuser detection iteration decoding algorithm. In addition, we give the definition of a AUER of $10^{-5}$, under the condition that $E_b/N_0$ is 0 dB and the length of transmit block is larger than a specified value, e.g., 784, verifying the validity of the proposed UDAS sets.

Index Terms—Multiple access, uniquely-decodable multi-amplitude sequence, adder channel, Shannon limit, active user detection.

I. INTRODUCTION

EXT generation multiple-access is expected to support massive users in the limited resources, and many works have been done on this topic [1], [2], [3]. In general, the multiple-access technique can be categorized into uncoordinated multiple-access and coordinated multiple-access [4]. The uncoordinated multiple-access is generally viewed as unsource multiple-access, in which each user shares the same transmission protocol without allocated signature by a base station (BS). Conversely, the coordinated scenario is to administer the users by a central processor, i.e., BS, and each user is assigned a 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], [8], [9], [10], [11], [12], which can support larger number of users than the classical orthogonal spreading codes. Besides the binary and ternary code sequences, Zadoff-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], [15], [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 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 multi-dimensional codebook. Due to the sparsity of multi-dimensional codebooks, the proposed scheme can maintain overloaded information and enable massive connectivity. When these classical pilot sequences (or multi-dimensional codebooks) are used as multiuser signatures for a GFMA system, there exist some challenges and/or drawbacks.

- The spectrum efficiency (SE) of a spreading sequence based multiple-access system is generally determined by the number of sequences and the length of a spreading sequence. For example, the SE of an orthogonal binary spreading sequence based multiple-access system is equal to one. When UD ternary codes are utilized, the SE can be larger than one, because of the overloaded information. It is a challenging research topic to design high SE UD ternary code sets for a MAC.

- 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], [20], [21], [22], [23]. Thus, the extension of the designed multiuser codebook 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], [25], [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 \( \{1, 2, \ldots, 2^{J-1}\} \), 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 (UDS or UDA sequence) for GFMA systems. A UDA sequence owns sequential multiple amplitudes, which can provide prior information for detection. The advantages of the proposed UDA mainly include three aspects. 1) It can support a high SE transmission without additional redundancy, which can compare with some designed non-orthogonal multiple-access (NOMA) codebooks. 2) It can be easily applied to various multiple-access scenarios with favored flexibility. More importantly, the proposed UDA can be further extended and applied to other scenarios, e.g., modulation, space-time block coding (STBC), multiple-input multiple-output (MIMO), and etc. 3) Based on the feature of UDA, the AUD algorithm can be easily realized with low-complexity.

This paper gives the definition of a \( T \)-size UDA set, and constructs two kinds of UDA sets based on cyclic and quasi-cyclic structures. The cyclic and quasi-cyclic structures can make the UDA sequences own the same power, which is helpful for realizing low-complexity AUD. Then, we propose an UDA-based multi-dimensional bit interleaving coded modulation (MD-BICM) transmitter and receiver, consisting of a statistic of UDA feature based AUD algorithm (SoF-AUD) and a multiuser detection (MUD) iteration decoding algorithm. The theoretical active user error rate (AUER) and the Shannon limits of the proposed system are also deduced in details.

The remainder of this paper is organized as follows. In Section II, we give some definitions of UDA, and construct two kinds of UDA sets. The UDA-based MD-BICM system is described in Section III. Section IV shows 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.

In this paper, \( a, a \) and \( A \) stand for a variable, a vector and a matrix, respectively. Denote \( A^T \) by the transpose of a matrix \( A \). 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 imaginary part of the complex number \( a \). \( i \) stands for an imaginary number. \( E[\cdot] \) is the function of expectation, and \( \lceil \cdot \rceil \) denotes the ceiling operator. \( \| a \|_2 \) denote the 2-norm of a vector \( a \).

II. DEFINITION, CONSTRUCTION, AND FEATURES OF UDA SETS

This section presents definition, construction and features of UDA in synchronous adder multiple-access channels. The well-behaved designed UDA sets can separate multiuser without ambiguity, and this paper introduces two of them.

A. Definition

According to [28], a multiuser uniquely-decodable mapping (UDM) element set is defined by \( \Delta = \{\Delta_{re}, \Delta_{im}\} \), where \( \Delta_{re} = \{1, 2, \ldots, 2^n\} \) and \( \Delta_{im} = \{1i, 2i, \ldots, 2pi\} \) with \( p \in \mathbb{Z} \). Denote \( |\Delta| \), \( |\Delta_{re}| \), and \( |\Delta_{im}| \) by the number of elements in sets \( \Delta, \Delta_{re}, \) and \( \Delta_{im} \), respectively. Obviously, it is derived that \( |\Delta_{re}| = |\Delta_{im}| = p + 1 \) and \( |\Delta| = 2(p + 1) \), which are determined by the parameter \( p \). The UDM element set can simultaneously support \( 2(p + 1) \) users without ambiguity, if each user is assigned different elements in \( \Delta \) and then utilizes the selected element and its inverse as its modulated symbols.

**Definition 1:** An \( L \)-length UDA sequence is consisted of \( L \) sequential elements, which are selected from an UDM element set \( \Delta \) with \( p \) as a parameter.

Let all the \( L \)-length UDA sequences 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 \( \Psi \) contains \( T \) sequences, as

\[
\Psi = \{e_{t,1}, e_{t,2}, \ldots, e_{t,l}, \ldots, e_{t,T}\},
\]

with \( e_{t,l} = (e_{t,1}, e_{t,2}, \ldots, e_{t,l}, \ldots, e_{t,T}) \in \mathbb{E}_L \), where \( e_{t,l} \in \Delta, 1 \leq t \leq T \) and \( 1 \leq l \leq L \). There are \( T \) users, and each user is assigned a sequence, i.e., \( e_{t,l} \), and the transmit bits of the \( t \)-th user is defined by \( e^{(t)} = (c^{(t)}_1, c^{(t)}_2, \ldots, c^{(t)}_l, \ldots, c^{(t)}_T) \) where \( c^{(t)}_l \in \mathbb{B}^{1 \times L} \) for \( 1 \leq l \leq T \). The modulated symbol of the \( t \)-th symbol of the \( l \)-th user is equal to \( (2c^{(t)}_l - 1) \cdot e_{t,l} \), thus the sum-pattern of \( T \) users \( w_t \) can be given as \( w_t = \sum_{l=1}^{\Omega} (2c^{(t)}_l - 1) \cdot e_{t,l} \), where \( w_t \in \Omega_T^L \) and \( |\Omega_T^L| = 2^T \) for \( 1 \leq l \leq L \). Then, the sum-pattern vector of the entire \( L \) symbols of the \( T \) users is defined by \( w = (w_1, w_2, \ldots, w_T, \ldots, w_L) \) that belongs to the sum-pattern set \( \Omega_T^L \), i.e., \( w \in \Omega_T^L \), where \( \Omega_T^L = \{\Omega_T^1, \Omega_T^2, \ldots, \Omega_T^L\} \). If the sequences in \( \Psi \) 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) It is a one-to-one mapping between the sum-pattern vector and the transmit vectors of $T$ users, i.e., $w \leftrightarrow \{c^{(1)}, c^{(2)}, \ldots, c^{(l)}, \ldots, c^{(T)}\}$; and

3) The power constraint satisfies

$$\lim_{L \to \infty} P( |P_l - P_{\text{avg}}| < \epsilon ) = 0,$$

where $P_l = \frac{1}{T} \sum_{i=1}^{T} P_l$ stands for the average power of the $t$th UDA set, $P_{\text{avg}} = \frac{1}{T} \sum_{i=1}^{T} P_l$ is the average power of the entire UDA set, and $\epsilon$ is a positive small number;

then, the set $\Psi$ is defined by a $T$-size UDA set.

The former two conditions are used to ensure that the superimposed signals of $T$ users can be uniquely separated. The last condition is used to keep fairness among users. We further analyze the three conditions one-by-one.

- The first condition is to guarantee that users can be uniquely separated without ambiguity at the location $l$. For example, assume $e_{l,T} = 1$ and $e_{l,T} = 1$, at this moment, we have $w_l \in \{-2,0,2\}$ which is not an UDM set. If $e_{l,T} = 1$ and $e_{l,T} = 2$, we have $w_l \in \{-3,-1,1,3\}$ which is an UDM set.

- The second condition is to maintain that the sum-pattern vector $w$ can be uniquely mapped to the transmit vectors of $T$ users. For example, assume $T = 2$, $L = 4$, $e_1 = (1,2,2,1)$ and $e_2 = (2,1,1,2)$. If $w = (3,-1,-3,-1)$, we can obtain $c^{(1)} = (1,-1,-1,1)$ and $c^{(2)} = (1,1,-1,-1)$.

- The third condition anticipates that 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, $\frac{\|w_l\|}{P_{\text{avg}}} < \epsilon$, which is used to make the peak-to-average power ratio (PAPR) of one user within an acceptable range.

Observe the definition of UDA set, it is found that $T$ and $L$ are two important parameters of an UDA set. $T$ stands for the number of UDA sequences, and $L$ is the length of an UDA sequence. In general, a large $T$ indicates that the UDA set can support more users, at the expense of a large $L$. Under power constraint, a large $|\Omega_l^{(T)}|$ may result in a reduced Euclidean distance of the sum-pattern of $T$ users, degrading the BER performance. Therefore, with the help of the mapping relationship between sum-pattern $w$ and $\{c^{(1)}, c^{(2)}, \ldots, c^{(T)}\}$, it is easy to realize multiuser detection (MUD). Normally, a UDA has two benefits, one is used as a signature; and the other is used for assisting MUD, since a UDA sequence can provide prior probability for MUD. There are three relationships between $T$ and $L$, they are $L = T$, $L < T$ and $L > T$. One of the special cases is that $L = 1$, at this moment, a UDA degrades to a simple UDM case.

This paper only introduces the conception of UDA set, which is a strict definition. It is appealing to extend the conception of UDA set to a general case, i.e., quasi UDA set. Due to the page limitation, more contents about UDA sets will be introduced in our future work.

B. Construction

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

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

$$E_{\text{cyc}} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_T \\ \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \ldots & a_T \\ a_T & a_1 & \ldots & a_{T-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_2 & a_3 & \ldots & a_1 \end{bmatrix}. \tag{1}$$

Because of the cyclic structure, the rows of $E_{\text{cyc}}$, i.e., $e_1, e_2, \ldots, e_T$, are formed a $T$-size UDA set $\Psi$. Moreover, it is easy to derive that $P_{\text{avg}} = P_l = \frac{1}{T} \sum_{i=1}^{T} P_l^2$, where $1 \leq t \leq T$ and $L = T$. We give an example to show the construct processing.

Example 1: Assume $T = 4$, and the generator $a = (1,1,2,2)$. Then,

$$E_{\text{cyc}}(4,4) = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ \end{bmatrix} = \begin{bmatrix} 1 & 1i & 2 & 2i \\ 2i & 1i & 1 & 2 \\ 2 & 2i & 1 & 1i \\ 1i & 2 & 2i & 1 \end{bmatrix}, \tag{2}$$

which is a 4-size UDA set, and the length of each sequence is 4. The average power $P_{\text{avg}}$ equals to 2.5. ▲▲

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

$$E_{\text{qc}} = \begin{bmatrix} A_{1,1} & A_{1,2} & \ldots & A_{1,Q} \\ A_{2,1} & A_{2,2} & \ldots & A_{2,Q} \\ \vdots & \vdots & \ddots & \vdots \\ A_{S,1} & A_{S,2} & \ldots & A_{S,Q} \end{bmatrix}, \tag{3}$$

where $A_{s,q}$ is an $L \times L$ cyclic matrix. Hence, $E_{\text{qc}}$ is an $(L \cdot S) \times (L \cdot Q)$ quasi-cyclic matrix. Denote the generator of $A_{s,q}$ by $a_{s,q}$, i.e., $a_{s,q} = (a_{s,q,1}, a_{s,q,2}, \ldots, a_{s,q,s}, \ldots, a_{s,q,Q})$, where $1 \leq s \leq S$, $1 \leq q \leq Q$ and $1 \leq l \leq L$. Given $q$ and $l$, if $s \neq s'$, $a_{s,q,l} \neq a_{s',q,l}$, then the rows of $E_{\text{qc}}$ can form an $(L \cdot S)$-size UDA set, in which each sequence is with length $L = L \cdot Q$.

Equation (3) is a general expression of the QC structure, which can be extended to a block-wise cyclic structure, as

$$E_{b,\text{qc}} = \begin{bmatrix} A_{1,1} & A_{1,2} & \ldots & A_{1,Q} \\ A_{1,Q} & A_{1,1} & \ldots & A_{1,Q-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1,2} & A_{1,3} & \ldots & A_{1,1} \end{bmatrix}, \tag{4}$$

where $A_{1,q}$ is an $L \times L$ 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 (4) has cyclic structure, thus $E_{b,qe}$ is an $(L \cdot Q) \times (L \cdot Q)$ block-wise cyclic square matrix, where “b” stands for “block-wise”.

**Example 2:** Let $Q = 3$, and the generators of $A_{1,1,1}, A_{1,2,1}$ and $A_{1,1,3}$ are respectively $a_{1,1} = (1, i, 1), a_{1,2} = (2, 2i)$ and $a_{1,3} = (4, 4i)$. Then, the block-wise QC structure $E_{b,qe}$ is given by

$$E_{b,qe}(6, 6) = \begin{bmatrix} 1 & 1i & 2 & 2i & 4 & 4i \\ 1i & 1 & 2i & 2i & 4 & 4i \\ 4i & 4i & 1 & 2i & 2 & 2i \\ 4i & 4i & 1 & 2i & 2 & 2i \\ 2 & 2i & 4i & 4i & 1 & 1i \\ 2i & 2i & 4i & 4i & 1 & 1i \end{bmatrix},$$

where $E_{b,qe}(6, 6)$ indicates a 6-size UDAS set, and the length of each sequence is equal to 6. Each row of $E_{b,qe}$ is an UDAS, with average power $P_{avg} = 7$. ▼▼

In fact, a long UDA sequence can be truncated into a short UDA sequence. Thus, it is easy to extend the constructed cyclic or quasi-cyclic matrix into a general case. For example, we can take a 6 × 4 sub-matrix of (5), and obtain

$$E_{b,qe}(6, 4) = \begin{bmatrix} 1 & 1i & 2 & 2i \\ 1i & 1 & 2i & 2 \\ 4i & 4i & 1 & 1i \\ 4i & 4i & 1 & 1i \\ 2 & 2i & 4i & 4 \\ 2i & 2i & 4i & 4 \end{bmatrix},$$

which is a 6-size UDAS set, and the length of each sequence is equal to 4.

**C. Features of the Proposed Cyclic/quasi-Cyclic UDAS Set**

As aforementioned definition, the sum-pattern vector set of $T$-size UDAS set $\Psi$ is $\Omega^T = (\Omega^T_1, \Omega^T_2, \ldots, \Omega^T_{T})$. However, for a random multiple access scenario, not all the $T$ users are simultaneously transmitted. Now, we consider a general case.

Assume $\tau \ (1 \leq \tau \leq T)$ different sequences of $\Psi$ are selected for multuser transmission, i.e., $\Psi^{(\tau)} = \{e_1, e_2, \ldots, e_{\tau}\}$, where $\tau \in \{1, 2, \ldots, T\}$ and $\mu$ is the $\tau$-size combination index. Actually, there are totally $C_T^\tau$ combinations, indicating $1 \leq \mu \leq C_T^\tau$. At this moment, the sum-pattern of the $\mu$th symbol of the $\mu$th combination index is $w_\mu = \sum_{\tau=1}^{\mu} \{2e^{(\mu)}_{\tau} - 1\} \cdot e_{\tau,1}$, where $w_\mu \in \omega_1^{(\tau)}$, and the number of elements in $\omega_1^{(\tau)}$ equals to $|\omega_1^{(\tau)}| = 2^\tau$.

Define the sum-pattern vector set of $\tau$ users by $\Omega^\tau = (\Omega^1_1, \Omega^1_2, \ldots, \Omega^1_T), (\Omega^2_1, \ldots, \Omega^1_T), \ldots, (\Omega^T_1, \ldots, \Omega^T_T)$, where $\Omega^\tau = (\omega_1^{(1)} \cup \omega_2^{(2)} \cup \ldots \cup \omega_t^{(t)} \cup \ldots \cup \omega_T^{(T)})$ and the number of elements in $\Omega^\tau$ satisfies $|\Omega^\tau| \leq C_T^\tau \cdot 2^\tau$. When $\tau = T$, it is found that $|\Omega^T| = 2^T$, equaling to the T-size case.

Moreover, define the maximum values of the real and imaginary parts of $\omega_1^{(\tau)}$ by

$$n_{r,\tau} = \max \{\text{Re}[\omega_1^{(\tau)}]\}, \quad n_{i,\tau} = \max \{\text{Im}[\omega_1^{(\tau)}]\}.$$  

While, $(k_{1,\tau, \tau}, n_{1,\tau, \tau})$ are used for finding $\omega_1^{(\tau, \mu)}$. The average power of the sum-patterns in $\omega_1^{(\tau, \mu)}$, defined by $\lambda_1^{(\tau, \mu)}$, is equal to

$$\lambda_1^{(\tau, \mu)} = \frac{1}{2^\tau} \sum_{\tau=1}^{\mu} |w_\mu|^2 = \sum_{\tau=1}^{\mu} 2^2 = \tau \cdot LP_{avg},$$

relying on the values of $\{e_{\tau,1}, e_{\tau,2}, \ldots, e_{\tau,T}\}$. For a given $\tau$, define the average power of $\omega_1^{(\tau, \mu)}$ of the total $L$ locations by $\Lambda^{(\tau, \mu)} = \{\lambda_1^{(\tau, \mu)}, \lambda_2^{(\tau, \mu)} \ldots, \lambda_T^{(\tau, \mu)}\}$, assisting AUD and multiuser detection. It is easy to derive that the sum of $\lambda_1^{(\tau, \mu)}$ is

$$\lambda_{sum} = \sum_{\tau=1}^{\mu} \lambda_1^{(\tau, \mu)} = \sum_{\tau=1}^{\mu} 2^2 = \tau \cdot LP_{avg},$$

which is a constant for a given $\tau$, independent of the selection of sequences, i.e., combination index $\mu$. 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 3:** Recall (2) of Example 1, there are four sequences, $\Psi = \{e_1, e_2, e_3, e_4\} = \{(1, 1i, 1, 2i), (1i, 1, 2, 1i), (2i, 1i, 1, 1i), (1i, 2i, 1i, 1i)\}$, with $P_{avg} = 2.5$ and $L \cdot P_{avg} = 10$.

If we select any $\tau = 2$ sequences from the given $\Psi$ 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 $\omega_1^{(\tau)}$ is equal to $|\omega_1^{(\tau, \mu)}| = 2^2 = 4$, where $1 \leq \mu \leq C_T^\tau$. When $l = 1$, we have

$$\{1, 2i\} \rightarrow \omega_1^{(2,1)} = \{1 + 2i, 1 - 2i, -1 + 2i, -1 - 2i\} \rightarrow \lambda_1^{(2,1)} = 5;$$

$$\{1, 2\} \rightarrow \omega_1^{(2,2)} = \{-3, -1, 1, 3\} \rightarrow \lambda_1^{(2,2)} = 5;$$

$$\{1, 1i\} \rightarrow \omega_1^{(2,3)} = \{1 + 1i, 1 - 1i, -1 + 1i, -1 - 1i\} \rightarrow \lambda_1^{(2,3)} = 2;$$

$$\{2i, 2\} \rightarrow \omega_1^{(2,4)} = \{2 + 2i, 2 - 2i, -2 + 2i, -2 - 2i\} \rightarrow \lambda_1^{(2,4)} = 8;$$

$$\{2i, 1i\} \rightarrow \omega_1^{(2,5)} = \{-3i, -1i, 1i, 3i\} \rightarrow \lambda_1^{(2,5)} = 5;$$

$$\{2, 1i\} \rightarrow \omega_1^{(2,6)} = \{2 + 1i, 2 - 1i, -2 + 1i, -2 - 1i\} \rightarrow \lambda_1^{(2,6)} = 5.$$

It is found that $\omega_1^{(\mu)} \cap \omega_1^{(\mu')}$ when $\mu \neq \mu'$. Moreover, it is able to derive that $\lambda_{sum} = 20$, which equals to $2 \times (L \cdot P_{avg})$.

Observe $\Lambda^{(2,2)}$, it is found that $\Lambda^{(2,2)} = \Lambda^{(5,2,5)}$, i.e., $(5, 5, 5, 5)$; thus, $\Psi^{(2,2)}$ and $\Psi^{(2,5)}$ cannot be separated based on their $\Lambda$. However, observe $\omega_1^{(2,2)}$ and $\omega_1^{(2,5)}$ again, it is found that $(k_{1,\tau, \tau}, k_{1,\tau, \tau}) = (3, 0)$ and $(k_{1,\tau, \tau}, k_{1,\tau, \tau}) = (0, 3)$, indicating that $\Psi^{(2,2)}$ and $\Psi^{(2,5)}$ can be separated based on their corresponding $(k_{1,\tau, \tau}, k_{1,\tau, \tau})$. 

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
If we pick any \( \tau = 3 \) sequences from \( \Psi \) for multiuser transmission, there are \( C_4^2 = 4 \) combinations, i.e., \( \Psi^{(3,1)} = \{e_1, e_2, e_3\}, \Psi^{(3,2)} = \{e_1, e_2, e_4\}, \Psi^{(3,3)} = \{e_1, e_3, e_4\}, \) and \( \Psi^{(3,4)} = \{e_2, e_3, e_4\} \). At this time, the sum-pattern sets of the \( \mu \)th combination of the \( l \)th symbol can be given. Take \( l = 1 \) as an example, they are

\[
\begin{align*}
\{1, 2i, 2\} \rightarrow & \omega_1^{(3,1)} = \{-3 + 2i, -3 - 2i, 3 + 2i, 3 - 2i, \}
\rightarrow \lambda_1^{(3,1)} = 9; \\
\{1, 2i, 1\} \rightarrow & \omega_1^{(3,2)} = \{-3i, -1 + 2i, -1 - 2i, -1 + i, -1 - i, 1 + i, -1 + 1i, 1 + 3i, -1 - 3i\}
\rightarrow \lambda_1^{(3,2)} = 6; \\
\{1, 2, l\} \rightarrow & \omega_1^{(3,3)} = \{-3 + 1i, -3 - 1i, 3 + 1i, 3 - 1i, \}
\rightarrow \lambda_1^{(3,3)} = 6; \\
\{2i, 2, 1\} \rightarrow & \omega_1^{(3,4)} = \{-2 + 3i, -2 - 3i, 2 - 1i, 2 - 2i, 2 + i, -2 + 1i, 2 + 3i, -2 + 3i\}
\rightarrow \lambda_1^{(3,4)} = 9.
\end{align*}
\]

For this scenario, it is found that \( \lambda_{\text{sum}}^{m=3} = 30 \) is equal to \( 3 \times (L \cdot P_{\text{avg}}) \). Moreover, we can see \( \omega_1^{(3,2)}(\omega_1^{(3,3)}) = \{1 - 1i, -1 - 1i, 1 - 1i, -1 - 1i\} \), indicating that there exists constellation overlap. Observe the four sets \( \Lambda^{(3,1)}, \Lambda^{(3,2)}, \Lambda^{(3,3)} \) and \( \Lambda^{(3,4)} \), it is a one-to-one mapping between \( \Lambda^{(3,\mu)} \) and \( \Psi^{(3,\mu)} \), i.e., \( \Lambda^{(3,\mu)} \rightarrow \Psi^{(3,\mu)} \). For this scenario, we can calculate \( \Lambda^{(3,\mu)} \), and then obtain the corresponding \( \Psi^{(3,\mu)} \).

### III. An UDAS-Based MD-BICM System

Suppose there are \( J \) users simultaneously access the BS. Each user is assigned a UDA sequence from the \( T \)-size UDAS set \( \Psi \), where the length of a UDA sequence in \( \Psi \) is equal to \( L \). In this paper, we assume that \( J \leq T \). Obviously, the value of \( J \) should be estimated for a GFMA system. The system model is shown in Fig. 2.

The transmit bit information of the \( j \)th user is defined by \( u_j^{(j)} \in \mathbb{B}_1^{1 \times K} \), where \( 1 \leq j \leq J \) and \( K \) is the number of bits of the \( j \)th user. \( u_j^{(j)} \) is then passed to the first channel encoder whose generator is a \( K \times N_1 \) matrix defined by \( G_1 = [g_{k,n_1}]_{1 \leq k \leq K, 1 \leq n_1 \leq N_1} \), and obtained an encoded codeword \( v_j^{(j)} = u_j^{(j)} \cdot G_1 \), where \( v_j^{(j)} = (v_1^{(j)}, v_2^{(j)}, \ldots, v_{N_1}^{(j)}) \).

We can write \( v_j^{(j)} \) into a matrix \( V_j^{(j)} \) column-by-column, given as

\[
V_j^{(j)} = \begin{bmatrix}
\begin{array}{cccc}
  v_1^{(j)} & v_{M+1}^{(j)} & \cdots & v_{(N_1-1)M+1}^{(j)} \\
  v_2^{(j)} & v_{M+2}^{(j)} & \cdots & v_{(N_1-1)M+2}^{(j)} \\
  \vdots & \vdots & \ddots & \vdots \\
  \end{array}
\end{bmatrix},
\]

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, \( V_j^{(j)} \) is utilized to show the interleaving process, since the encoded codeword \( v_j^{(j)} \) will be further processed row-by-row. Note that the number of columns of \( V_j^{(j)} \) is related to the length of the UDA sequence, i.e., \( L \).

Then, we equidistantly select \( L - 1 \) bits from each row of \( V_j^{(j)} \) where \( L \leq N_c \), and then pass the selected \( L - 1 \) bits to the second encoder \( G_2, G_3 \) can be as simple as possible, e.g., single parity-check code (SPC). For the \( m \)th row of \( V_j^{(j)} \) for \( 1 \leq m \leq M \), let the parity-check bit generated by \( G_2 \) be \( v_{m,\text{pc}}^{(j)} \), given by

\[
v_{m,\text{pc}}^{(j)} = v_{m+M \cdot (\log_2 M-1)}^{(j)} \oplus v_{m+M \cdot (2 \log_2 M-1)}^{(j)} \oplus \cdots \times \oplus v_{m+M \cdot ((L-1) \log_2 M-1)}^{(j)},
\]

which is located at the end of the \( m \)th row of \( V_j^{(j)} \).
Thereafter, we can achieve an $M \times N$ matrix $C^{(j)}$, as

$$C^{(j)} = \begin{bmatrix} c^{(j)}_1 \\ \vdots \\ c^{(j)}_M \end{bmatrix},$$

where

$$c^{(j)}_m = \left( c^{(j)}_{m,1}, \ldots, c^{(j)}_{m,l}, \ldots, c^{(j)}_{m,L} \right)$$

is a $1 \times N$ vector, with $N = N_c + 1$.

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

$$c^{(j)}_m = \left( c^{(j)}_{m,1}, \ldots, c^{(j)}_{m,l}, \ldots, c^{(j)}_{m,L} \right)$$

where $c^{(j)}_{m,l,b} \in \mathbb{B}$ for $1 \leq m \leq M$, $1 \leq l \leq L$, and $1 \leq b \leq \log_2 M$, and $c^{(j)}_{m,l} = (c^{(j)}_{m,1}, c^{(j)}_{m,2}, \ldots, c^{(j)}_{m,l,\log_2 M})$.

Look back to the parity-check bit $v^{(j)}_{m,pc}$, which can also be given as

$$v^{(j)}_{m,pc} = c^{(j)}_{m,1,\log_2 M} + c^{(j)}_{m,2,\log_2 M} + \cdots + c^{(j)}_{m,L,\log_2 M}.$$

Then, every $\log_2 M$ bits of $c^{(j)}_m$ are modulated to one symbol, i.e., $c^{(j)}_{m,l} \rightarrow x^{(j)}_{m,l}$. We can achieve a modulated $M \times L$ matrix $X^{(j)}$ as

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

where $x^{(j)}_{m,l}$ is an $M_1$-dimensional modulated symbol with the assumption of $M = 2M_1$, and belongs to one of the following forms,

$$s^{(j)}_{m,l} = (s^{(j)}_{m,l,0}, \ldots, s^{(j)}_{m,l,L}),$$

in which only one position of $x^{(j)}_{m,l}$ has value $s^{(j)}_{m,l}$ and the other $M_1 - 1$ positions are all zeros, for $1 \leq m \leq M$ and $1 \leq l \leq L$.

Note that $M_1$-dimensional modulation can increase both the flexibility and the minimum Euclidean distance, thus it is appealing for GFMA.

If the $i$th location of $x^{(j)}_{m,l}$ is non-zero, we set the location index $p^{(j)}_{m,l,i} = 1$, and other location indexes $p^{(j)}_{m,l,i'} = 0$ for $i' \neq i$ and $1 \leq i' \leq M_1$. The non-zero location is determined by the first $\log_2 M_1$ bits of $c^{(j)}_{m,l}$.

Suppose the $j$th user selects $e_j = (e_{j,1}, e_{j,2}, \ldots, e_{j,L})$ as its signature, then $s^{(j)}_{m,l}$ is calculated as

$$s^{(j)}_{m,l} = \left( 2^{e_{j,1,\log_2 M_1}} - 1 \right) \cdot e_{j,l},$$

which is determined by the last bit $c^{(j)}_{m,l,\log_2 M} = c^{(j)}_{m,l}$ of $c^{(j)}_{m,l}$ and the $i$th symbol of $e_j$.

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

**Example 4**: Assume a $1 \times 21$ vector $V^{(j)}$ is given as $V^{(j)} = (0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0)$, so that $N_1 = 21$. Set $M = 2$, $N_c = 11$, $M_1 = 8$ and $M_1 = 4$, then $V^{(j)}$’s corresponding matrix $V^{(j)}$ is given by

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

and $v_{21,31}$ is set to be 0 (shown in red color). Then, SPC is done to each row of $V^{(j)}$, given by $v^{(j)}_{1,p_1} = v_5 \oplus v_{11} \oplus v_{17} = 0$ and $v^{(j)}_{2,p_1} = v_6 \oplus v_{12} \oplus v_{18} = 1$. Thus, it is obtained

$$C^{(j)} = \begin{bmatrix} c^{(j)}_1 \\ c^{(j)}_2 \end{bmatrix} = \begin{bmatrix} 011 & 100 & 001 & 110 \\ 100 & 010 & 111 & 001 \end{bmatrix},$$

which is a $2 \times 12$ matrix, where $N = 12$ and $L = 4$.

Assume ‘00’ , ‘01’, ‘10’ and ‘11’ are respectively mapped to four different locations, suppose $e_j = (1, 1, 2, 2i)$, it is able to obtain

$$x^{(j)}_{1,1} = (0, 0, 0, 0), \quad x^{(j)}_{1,2} = (0, 0, -1i, 0),$$

$$x^{(j)}_{1,3} = (2, 0, 0, 0), \quad x^{(j)}_{1,4} = (0, 0, 0, -2i),$$

$$x^{(j)}_{2,1} = (0, 0, -1, 0), \quad x^{(j)}_{2,2} = (0, -1i, 0, 0),$$

$$x^{(j)}_{2,3} = (0, 0, 0, 0), \quad x^{(j)}_{2,4} = (0, 0, 0, 0).$$
Then \( x_{2,3}^{(j)} = (0, 0, 0, +2) \), \( x_{2,4}^{(j)} = (+2i, 0, 0, 0) \).

Then, \( x_{1}^{(j)} \) and \( x_{2}^{(j)} \) are transmitted to the channel. ▲▲

Then, \( X(j) \) is sent to the adder multiple-access channel row-by-row. At the receiver, the received signal of the \( m \)th row is equal to

\[
y_m = \sum_{j=1}^{J} x_m^{(j)} + z_m,
\]

where \( 1 \leq m \leq M \), and \( z_m \) is an additive white Gaussian noise (AWGN) vector. Each noise component of \( z_m \) is an independent and identically distributed \((i.i.d)\) Gaussian random variable with distribution \( N(0, N_0/2) \). Set \( y = (y_1, y_2, \ldots, y_M) \), where \( y_m = (y_{m,1}, y_{m,2}, \ldots, y_{m,l}, \ldots, y_{m,L}) \) includes \( L \) symbols. For the \( l \)th received symbol \( y_{m,l} \) of \( y_m \), we have \( y_{m,l} = (y_{m,1}, y_{m,2}, \ldots, y_{m,l}, \ldots, y_{m,L}) \), with

\[
y_{m,l,i} = \sum_{j=1}^{J} p_{m,l,i}^{(j)} x_{m,l}^{(j)} + z_{m,l,i},
\]

where \( p_{m,l,i}^{(j)} = \{0, 1\} \) and \( 1 \leq i \leq M_1 \). Based on the received \( y \), the AUD modulator can detect the number of arrival users and identify users’ signatures. Afterwards, by a MUD detection algorithm, all the users’ information bits can be recovered.

In summary, the first channel encoder is mainly used for improving the reliability of the system, and the second channel encoder is used for assistant AUD. Since the second encoder is a simple SPC, it can realize hard decision for quick AUD with an acceptable detection performance. Note that both of the encoders \((G_1 \text{ and } G_2)\) can be removed, and the system still works, which reflects the great flexibility of the proposed scheme. In addition, \( M_1 \)-dimensional modulation can increase both the flexibility and the minimum Euclidean distance, at the expense of the increased spectral bandwidth. Therefore, we can set \( M_1 = 1 \) for achieving a higher SE, and set \( M_1 > 1 \) for improving the BER performance. More importantly, some coded schemes can be further applied to the \( M_1 \) orthogonal resources.

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 \( M_1 \)-dimension modulation, where \( 1 \leq j \leq J, 1 \leq m \leq M, 1 \leq l \leq L, 1 \leq b \leq \log_2 M \) and \( 1 \leq i \leq M_1 \).

IV. ACTIVE USER DETECTION AND MULTIUSER DETECTION ALGORITHMS

This section presents a statistic of UDAS feature based active user detection (SoF-AUD), and a multiuser detection (MUD) iteration decoding algorithm for the proposed system. Assume that the users are assigned different sequences from the \( T \)-size UDAS set \( \Psi \). The receiver will detect the number of arrival users and recover data sequences for all the users.

A. MAP Detection

First of all, we introduce the maximum a posterior (MAP) detection algorithm, which is an optimal solution for the proposed system. Take \( (6) \) into \( (8) \), \( y_{m,l,i} \) is rewritten as

\[
y_{m,l,i} = \sum_{j=1}^{J} p_{m,l,i}^{(j)} \cdot \left( 2c_{m,l,\log_2 M} - 1 \right) \cdot e_{j,l} + z_{m,l,i},
\]

where \( w_{m,l,i} = \sum_{j=1}^{J} p_{m,l,i}^{(j)} \cdot \left( 2c_{m,l,\log_2 M} - 1 \right) \cdot e_{j,l} \) is the received superimposed signal (or called sum-pattern). For further discussion, set \( w = (w_1, w_2, \ldots, w_m, \ldots, w_M) \), and \( w_m = (w_{m,1}, w_{m,2}, \ldots, w_{m,l}, \ldots, w_{m,L}) \), where \( w_{m,l} = (w_{m,1}, w_{m,2}, \ldots, w_{m,l}, \ldots, w_{m,L}) \).

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

\[
\left( J, C, \Psi^{(J,\mu)} \right) = \arg \max \left( P_J \cdot \exp \left\{ -\frac{\|y - w\|^2}{N_0} \right\} \right),
\]

where \( C = \{C^{(1)}, \ldots, C^{(J)}\} \) 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 \( T \) users. When \( J > T \), the detection is interrupted and all the users’ packets are lost. To accord with the aforementioned definition on UDAS set, in the following discussion, we set \( J = \tau \).

Although the MAP detection is the optimal algorithm, the complexity is extremely high for the proposed system, i.e., \( O(M^\tau L) \), which is exponential increased with the parameters of \( M \), \( L \) and \( T \). For example, when \( M = 2, L = 2 \) and \( T = 2 \), it is about \( 2.58 \times 10^{12} \), which is intolerable. Thereby, we present a low complex SoF-AUD algorithm, followed by a MUD algorithm.

B. Low-Complexity SoF-AUD and MUD Algorithms

At the receiver, it should detect the number of arrival users, separate the superimposed signals, and recover the transmit signals of all the users. We will introduce them step-by-step.

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

Define the \( i \)th location of the \( l \)th received symbol of all \( M \) rows by

\[
y_{l,i}^{\text{loc}} = (y_{1,l,i}, y_{2,l,i}, \ldots, y_{m,l,i}, \ldots, y_{M,l,i}).
\]

If we ignore the effect of noise and \( M \) is large enough, \( y_{l,i}^{\text{loc}} \) can traverse all the constellations of \( \omega_{1}^{(\tau,\mu)} \), including \( (\kappa_{1,l,i}^{(\tau,\mu)}, \kappa_{1,l,i}^{(\tau,\mu)}) \). Actually, the proposed SoF-AUD algorithm relies on the statistical properties of \( y_{l,i}^{\text{loc}} \).

The sum power \( \gamma_{\text{sum}} \) of the entire \( L \) symbols can be calculated as

\[
\gamma_{\text{sum}} = \sum_{l=1}^{L} \gamma_l = \sum_{l=1}^{L} \left\{ \frac{1}{M} \sum_{m=1}^{M} \sum_{i=1}^{M_1} \|y_{m,l,i}\|^2 \right\},
\]

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
where $\gamma_l = \frac{1}{M} \sum_{m=1}^{M-1} \left[ \sum_{i=1}^{M_1} \| y_{m,l,i} \|^2 \right]$ is the average power of the $l$th symbol.

Thus, for a given number of arrival users $\tau$, the statistical mean of $\gamma_{\text{sum}}$ can be deduced as

$$\tau_{\text{sum}} = \mathbb{E} \left[ \sum_{l=1}^{L} \left\{ \frac{1}{M} \sum_{m=1}^{M_1} \| y_{m,l,i} \|^2 \right\} \right]$$

$$= \mathbb{E} \left[ \sum_{l=1}^{L} \sum_{m=1}^{M_1} \| w_{m,l,i} + z_{m,l,i} \|^2 \right]$$

$$= \sum_{l=1}^{L} \sum_{m=1}^{M_1} \mathbb{E} \left[ \| w_{m,l,i} + z_{m,l,i} \|^2 \right]$$

$$= \lambda_1 + L \lambda_1 \cdot N_0, \quad (12)$$

since $\mathbb{E} [w_{m,l,i} + z_{m,l,i}] = 0$ and $\mathbb{E} [\| w_{m,l,i} \|^2] = N_0$. Equation (12) is the sum of $\lambda_1$ and $L \lambda_1 \cdot N_0$, indicating $\tau_{\text{sum}}$ is a constant for a given $\tau$. Here, $\lambda_1 = \tau \cdot (LP_{\text{avg}})$, and $LP_{\text{avg}}$ is a constant for a given cyclic/quasi-cyclic UDAS set $\Psi$.

Consider $M_1$-dimensional modulation, the maximum power of the $l$th symbol in $y_{\text{loc},i}$ can be calculated as

$$\zeta_{l,rc} = \max_{1 \leq m \leq M} \left\{ \text{Re} [y_{\text{loc},i}] \right\},$$

$$\zeta_{l,im} = \max_{1 \leq m \leq M} \left\{ \text{Im} [y_{\text{loc},i}] \right\}.$$ (13)

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

$$\left( \hat{\tau}, \hat{\mu} \right) = \arg \min_{1 \leq l \leq L} \| \gamma_{\text{sum}} - \tau_{\text{sum}} \|$$

$$= \arg \min_{1 \leq l \leq L} \| \gamma_{\text{sum}} - (\lambda_1 + L \lambda_1 \cdot N_0) \|,$$

s.t. $C1 : \hat{\mu} = \arg \min_{\mu} \sum_{l=1}^{L} d_{\tau}^{(\tau,\mu)},$ (14)

where $d_{\tau}^{(\tau,\mu)} = \| \zeta_{l,rc} - \zeta_{l,rc}^{(\tau,\mu)} \|^2 + \| \zeta_{l,im} - \zeta_{l,im}^{(\tau,\mu)} \|^2.$

C1 is used to find the uniquely selected UDAS set, more than one $\hat{\mu}$ may satisfy (13). Then, we can obtain the number of arrival users $\hat{\tau}$ and the selected UDAS set $\Psi(\hat{\tau},\hat{\mu})$.

Actually, the SoF-AUD is a kind of energy detection algorithm, thus its complexity is mainly determined by the sum of all the received signals’ power. According to (11), it is derived that the complexity of the proposed SoF-AUD algorithm is approximately $O(L M_1 M)$, which is linear increased with the parameters $L$, $M$, and $M_1$.

2) Step 2: MSE Detection: Based on the obtained $\hat{\tau}$ and $\Psi(\hat{\tau},\hat{\mu})$ by the SoF-AUD step, we begin to separate the superimposed signals. Since the statistics of all the received signals, i.e., $y_1, y_2, \ldots, y_M$, have been taken into consideration during the AUD process, we can deal with the received signals row-by-row to reduce the complexity.

When $M_1 = 1$, $s_{\text{loc}}^{(\tau)}$ is equivalent to $s_{\text{loc}}^{(\tau)}$. The detected results $\hat{\omega}_{m,1}$ can be achieved as

$$\hat{\omega}_{m,1} = \arg \min_{\omega_{m,1}} \| y_{m,1} - \omega_{m,1} \|^2,$$

s.t. $C2 : \Psi(\hat{\tau},\hat{\mu}) \rightarrow \hat{\omega}_{m,1}$

$$\rightarrow \{ c_{m,1,1}, c_{m,1,2}, \ldots, c_{m,1,L} \};$$

$C3 : 0 = c_{m,1,1}^{(j)} + c_{m,1,2}^{(j)} + \ldots + c_{m,1,L}^{(j)}.$ (14)

$C2$ stands for the one-to-one mapping between $\hat{\omega}_{m,1}$ and $\{ c_{m,1,1}, c_{m,1,2}, \ldots, c_{m,1,L} \}$. $C3$ reflects the SPC constraint of $\hat{\tau}$, where $1 \leq j \leq \hat{\tau}$. It is noted that (14) is a hard decision processing, which helps fast realize MUD with acceptable BER. It can be applied to the scenario without $G_1$.

When $M_1 > 1$, the detection processes becomes a little complex. Define the sum-pattern set of the $i$th location of the $M_1$-dimensional modulation by $\psi(\hat{\tau},\hat{\mu})$, where $\hat{\tau}_i$ and $\hat{\mu}_i$ are respectively $\hat{\tau}$ users and the $\hat{\mu}_i$th combination at the $i$th location. Let the set of $M_1$-dimensional sum-pattern be $\Theta(\hat{\tau},\hat{\mu}) = (\psi(\hat{\tau},\hat{\mu}_1), \psi(\hat{\tau},\hat{\mu}_2), \ldots, \psi(\hat{\tau},\hat{\mu}_L), \ldots, \psi(\hat{\tau},\hat{\mu}_L))$, the detected results are then expressed as

$$\hat{\omega}_{m,l} = \arg \min_{\theta_{m,l} \in \Theta(\hat{\tau},\hat{\mu})} \| y_{m,l} - \theta_{m,l} \|,$$

s.t. $C4 : \hat{\tau} = \sum_{i=1}^{M_1} \hat{\tau}_i,$

$C5 : \mu_i \leq \mu,$

$C6 : \hat{\omega}_{m,l} \rightarrow \{ c_{m,1,1}, c_{m,1,2}, \ldots, c_{m,1,L} \},$

$C7 : 0 = c_{m,1,1}^{(j)} + c_{m,1,2}^{(j)} + \ldots + c_{m,1,L}^{(j)}.$ (15)

where $\hat{\omega}_{m,l} = (\hat{\omega}_{m,1,l}, \hat{\omega}_{m,1,2}, \ldots, \hat{\omega}_{m,1,M_1}), \theta_{m,l} = (\omega_{m,1,l}, \omega_{m,1,2}, \ldots, \omega_{m,1,M_1})$ and $\omega_{m,1,i} \in \Psi(\hat{\tau},\hat{\mu}_i)$ for $1 \leq i \leq M_1$. $C4$ 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 $\hat{\omega}_{m,l}$ and $\{ c_{m,1,1}, c_{m,1,2}, \ldots, c_{m,1,L} \}$. $C7$ also stands for the SPC constraint of $G_2$, where $1 \leq j \leq \hat{\tau}$. While, if $G_2$ is not used in the system, we can ignore the $C7$ constraint. Essentially, (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 $G(\tau,\mu)$.

It is able to derive that the complexity of the MSE detection is approximately $O(ML \cdot \hat{\tau}^2)$, which is linear increased with the parameters $L$, $M$ and the number of superposition constellations $M^2$. Since there are totally $ML$ symbols, we only need to compare each received symbol with the number of superposition constellations $M^2$. Thus, the complexity is significantly reduced.

3) Step 3: Message Passing Iteration: When a LDPC code is utilized as the first encoder $G_1$, it is important to find the initial LLR (log likelihood ratio) according to the received signals for further decoding. In other words, we should obtain the initial LLRs from the MSE detection step.

The probability of $\hat{\omega}_{m,l}$ can be calculated as

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

where $\hat{\omega}_{m,l} \in \Theta(\hat{\tau},\hat{\mu})$. Since $\hat{\omega}_{m,l}$ is corresponding to log $2^M$ bits of $\hat{\tau}$ users, i.e.,
The MSE detection step. In other words, the MUD iteration takes into consideration the joint Tanner graph \( G \) by some of the variable nodes (VNs) of encoders be \( m, l \to j \) (\( j \)).

In (14), we can achieve the soft LLRs of \( c_{m,l,b}^{(j)} \) and CNs of \( G \) column-by-column, we can achieve the soft LLRs of \( v^{(j)} \) that are used for further decoding.

Assume the parity-check matrix of the first encoder is \( H_1 \). Since the second encoder utilizes SPC, encoders one and two can be jointly decoding. Let the Tanner graphs of the two encoders be \( G_1 \) and \( G_2 \), and the two graphs are connected by some of the variable nodes (VNs) of \( G_1 \) and check nodes (CNs) of \( G_2 \). Thus, the VNs update and CNs update should take into consideration the joint Tanner graph \( G \), as shown in Fig. 3. The proposed decoding algorithm is based on the joint Tanner graph \( G \), and the major alternations are the VNs of \( G_1 \) update and CNs of \( G_2 \) update.

- VNs of \( G_1 \) update. Since some of the VNs of \( G_1 \) are connected by both the CNs of \( G_1 \) and CNs of \( G_2 \), the update LLRs of these VNs should taken into consideration both CNs set.
- CNs of \( G_2 \) update. For the CNs of \( G_2 \), they are connected by both its own VNs and some VNs of \( G_1 \), thus, the update LLRs of the CNs of \( G_2 \) are determined by the two VNs sets.

It is noted that the initial LLRs of MPA are decided by the MSE detection step. In other words, the MUD iteration decoding includes both MSE detection and message passing iteration, which cooperatively realizes MUD of the proposed system. The entire receiver detection algorithm is summarized in Algorithm 1. In the Algorithm 1, the VNs and CNs of \( G_1 \) are defined by \( \alpha_1 \) and \( \beta_1 \); similarly, the VNs and CNs of \( G_2 \) are set to be \( \alpha_2 \) and \( \beta_2 \). Define \( N(\alpha_1), N(\alpha_2), N(\beta_1) \) and \( N(\beta_2) \) by the sets that are connected with \( \alpha_1, \alpha_2, \beta_1 \) and \( \beta_2 \), respectively. Note that, there is no edge between \( \alpha_2 \) and \( \beta_1 \), because of the SPC structure.

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

\[
P_{e,AU} = \Pr[\tau \neq \tilde{\tau}],
\]

where \( \tau \) and \( \tilde{\tau} \) 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 \( \gamma_{sum} \), which is affected by several parameters that are the sum-pattern set of \( \tau \) users \( \Omega^\tau \), the number of rows \( M \),
and the multi-dimensional order $M_1$. Thus, the AUER can be derived as
\begin{equation}
P_{c,AU} = \frac{1}{\tau-1} \sum_{\tau=1}^{T} \sum_{1 \leq \tau' < \tau, \tau' \neq \tau} \int_{D_{\tau'}} p(\gamma_{\text{sum}} | \tau) d\gamma_{\text{sum}},
\end{equation}
where $P_\tau$ is a priori probability of the number of arrival users, $p(\gamma_{\text{sum}} | \tau)$ is the conditional probability density function (PDF) of $\gamma_{\text{sum}}$ given by the number of arrival users $\tau$ for $\tau = 1, 2, \ldots, T$, and $D_{\tau'}$ is the decision region of $\tau'$.

Thereafter, the key issue is to derive the PDF of $p(\gamma_{\text{sum}} | \tau)$. Regarding as the sum-form of $\gamma_{\text{sum}}$ as shown in (11), $p(\gamma_{\text{sum}} | \tau)$ is determined by a sequence of noncentral Chi-Square random variables, defined by
\begin{equation}
p(\gamma_{\text{sum}} | \tau) = \sum_{s_a \in S} P(s_a) \cdot f_{\chi^2}(s_a),
\end{equation}
where $f_{\chi^2}(s_a)$ is the PDF of a noncentral Chi-Square random variable $\gamma_s$, where $\gamma_s = M \cdot \gamma_{\text{sum}}$, given by
\begin{equation}
f_{\chi^2}(s_a) = \frac{1}{N_0} \left( \frac{\gamma_s}{N_0} \right)^{N_s/2 - 1} e^{-\frac{\gamma_s}{N_0}} \cdot I_{N_s/2 - 1} \left( \frac{s_a}{N_0/2} \sqrt{\gamma_s} \right),
\end{equation}
where $I_n(r) = \sum_{k=0}^{\infty} \frac{(r/2)^{n+2k}}{2^{n+k} k! (n+k)!}$ and $\Gamma(r) = \int_0^\infty t^{r-1} e^{-t} dt$. $\gamma_s$ is consisted of $N_f$ independent Gaussian variables with comment variance $N_0/2$ and different means denoted by $s_{n_f}$. Since $r_m,i,i$ 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 $S$, i.e., $s_a \in S$. The probability of $s_a$ in set $S$ is defined by $P(s_a)$, thus, $\sum_{s_a \in S} P(s_a) = 1$.

Then, analyzing $p(\gamma_{\text{sum}} | \tau)$ is equivalent to finding the values of $s_a$ and $P(s_a)$, which can be calculated by the sum-pattern set $\Omega$. Let us review the expression of $\Omega$, which is defined by $\Omega = (\Omega_1^\tau, \Omega_2^\tau, \ldots, \Omega_t^\tau, \Omega_{t+1}^\tau, \ldots, \Omega_{\tau}^\tau)$ with $\Omega_i^\tau = \omega(\tau,1) \cup \ldots \cup \omega(\tau, \mu)$.

Divide the complex number set $\omega(\tau, \mu)$ into two sets, the real number set $\omega(\tau, \mu)$ and the imaginary number set $\omega_{\tau, \mu}$. Evidently, $s_{n_f}$ belongs to the set $\omega(\tau, \mu)$, i.e., $s_{n_f} \in \omega(\tau, \mu)$, varying with the varied $\tau$ and $\mu$, so to the variable $s_a$. For example, if $\tau = 1$, it is known that $s_a = \sqrt{\sum_{n_f=1}^{N_f} s_{n_f}^2} = \sqrt{M \sum_{l=1}^{L} \alpha_l^2} = \sqrt{M \cdot P_{\text{avg}}}$. In general, $s_a$ and $P(s_a)$ are both affected by the ratio of the number of elements in $\omega(\tau, \mu)$ to the total number of elements in $\Omega_{\tau}^\tau$. Therefore, different UDAS sets have different $s_a$ and $P(s_a)$, resulting in different $p(\gamma_{\text{sum}} | \tau)$.

To explain the calculation of $p(\gamma_{\text{sum}} | \tau)$, we present an example. Assume we utilize the UDAS set given by Example 2 as shown in (5), then
\begin{align*}
p(\gamma_{\text{sum}} | \tau = 1) &= f_{\chi^2}(\sqrt{2L \cdot P_{\text{avg}}}), \\
p(\gamma_{\text{sum}} | \tau = 2) &= \frac{2}{3} f_{\chi^2}(\sqrt{20M}) \\
&+ \frac{1}{3} \sum_{k=0}^{L_M} p(k, N_f) f_{\chi^2}(\sqrt{9k + (L_M - k)}),
\end{align*}
where $p(k, N_f) = C_{N_f}^{k} \cdot \left(\frac{1}{2}\right)^{N_f}$ is the probability of a binomial distribution. The case of $\tau = 1$ is accord with the aforementioned discussion.

When $\tau = 2$, $\mu$ belongs to $[1, C_T]$ and $\Psi(2, \mu)$ has been given in Example 4. Specifically, they can be divided into the following two cases.

1) Case 1: $\mu = 1, 3, 4$ or 6. For this case, the sum-patterns of the two users include both real and imaginary numbers. Take $\mu = 1$ as an example, we can get $\omega(1, 2) = \{1 + 2i, 1 - 2i, -1 + 2i, -1 - 2i\}$, $\omega(2, 1) = \{1 + i, 1 - i, -1 + i, -1 - i\}$, $\omega(2, 3) = \{2 + i, 2 - i, -2 + i, -2 - i\}$ and $\omega(2, 4) = \{2 + 2i, 2 - 2i, -2 + 2i, -2 - 2i\}$. 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 $s_a = \sqrt{M \cdot (5 + 2 + 5 + 8)} = \sqrt{20M}$, so do the cases of $\mu = 3, 4$ and 6.

2) Case 2: $\mu = 2$ or 5. At the moment, the sum-patterns of the two users are all real numbers (or imaginary numbers). Obviously, we can get $\omega(2, 2) = \{+3i, +i, -1i, -3i\}$ and $\omega(2, 5) = \{+3i, +i, -1i, -3i\}$. 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 $N(0, N_0/2)$, and the other $LM$ variables are in probability distributed as $N(\pm3, N_0/2)$ or $N(\pm1, N_0/2)$. Assume $k$ Gaussian variables are distributed as $N(\pm3, N_0/2)$, and the rest $LM - k$ Gaussian variables are distributing as $N(\pm1, N_0/2)$, where $0 \leq k \leq LM$ with probability $C_{LM}^{k} \cdot \left(\frac{1}{2}\right)^{LM}$. 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 $p(\gamma_{\text{sum}} | \tau = 2)$ 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 noted that the detection error probability of the UDAS set $\Psi(\tau, \mu)$ can be ignored compared to the AUER.

**B. Shannon Limits**

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

If all the users hold the same data rate, i.e., \( R_1 = R_2 = \ldots = R_J = R \), the last inequality dominates the others; while, the Shannon limit of the proposed system is deduced based on this assumption. Besides, due to the cyclic UDAS set, it is known that \( P_1 = P_2 = \ldots = P_J = P_{\text{avg}} \). Thus, the relationship between \( P_{\text{avg}} \) and \( L_b \) is 
\[ P_{\text{avg}} = \frac{R_{\text{avg}}(\lambda_1, M_1)}{T_s}, \]  
where \( T_s \) is the symbol duration.

With this assumption, the SE of the proposed system, defined by \( \lambda \), can be given as
\[ \lambda = \frac{J \cdot R_c \cdot \log_2 M}{M_1}, \]  
(25)
where \( M_1 \) reflects the number of resources caused by the multi-dimensional modulation. If \( M_1 = 1 \), then \( M = 2 \). Equation (25) becomes as
\[ \lambda = J \cdot R_c \leq T \cdot R_c, \]
due to the assumption of \( J \leq T \). Therefore, the SE is improved with the increased \( T \). For example, if \( T = 4 \) and \( R_c = 1 \), the SE of the proposed system equals to \( \lambda = 4 \).

It is noted that the length of a UDA sequence does not affect the SE, since the information bits are only modulated at the front signs of the UDAS. For example, if the front sign is “+”, it means that the transmit bit is “1”; otherwise, if the front sign is “−”, it indicates that the transmit bit is “0”.

In the following discussion, assume that the number of arrival users \( J = \tau \) and the selected UDAS set \( \Psi(\tau, \mu) \) have been available. Denote \( \{ x_{l_{m,i}}^{(j)} \}_{1 \leq m \leq M} \) by the transmit constellation set of the \( j \)th symbol of the \( j \)th user, including \( M \) possible \( M_1 \)-dimension transmit signals, expressed as 
\[ x_{l_{m,i}}^{(j)} = \{ x_{l_{m_i},1}^{(j)}, x_{l_{m_i},2}^{(j)}, \ldots, x_{l_{m_i},M_1}^{(j)} \}, \]
where \( x_{l_{m_i},i} \in \{ -e_{t_i}, 0, e_{t_i} \} \). With the assumption that the \( j \)th user selects the UDAS \( e_{t_j} \) in \( \Psi(\tau, \mu) \), Define \[ \Pr(x_{l_{m,i}}^{(j)}) \] by the prior probability of \( x_{l_{m,i}}^{(j)} \).

When \( J \) users’ signals, i.e., \( x_{l_{m,i}}^{(j)} \) for \( 1 \leq j \leq J \), are transmitted through a MAC, the PDF of the \( i \)th symbol of the received signal \( y_{m,l} \) is
\[ P(y_{m,l}|x_{l_{m,i}}^{(1)}, x_{l_{m_i}}^{(2)}, \ldots, x_{l_{m_i}}^{(j)}) = \prod_{m=1}^{M} \prod_{m=1}^{M} \frac{1}{\sqrt{\pi N_0}} \exp \left\{ -\frac{(y_{m,l} - \sum_{j=1}^{J} x_{l_{m,i}}^{(j)})^2}{N_0} \right\}. \]  
(26)
Note that the row index has no effect to the PDF. Consider the transmission of each user, the channel capacity of the $i$th symbol in the adder MAC can be calculated as

$$C_i = \max I(Y; X_1, X_2, \ldots, X_J)$$

$$= \max \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(y_{m,1}, x_{1,m}^{(1)}, x_{1,m}^{(2)}, \ldots, x_{1,m}^{(J)}) \cdot \log_2 \left[ \frac{p(y_{m,1}, x_{1,m}^{(1)}, x_{1,m}^{(2)}, \ldots, x_{1,m}^{(J)})}{p(y_{m,1})} \right] \, dy_{m,1}. \quad (27)$$

where $Y$ is the received signal’s variable, $X_1, X_2, \ldots, X_J$ are the transmit signals’ variables of $J$ users.

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

$$C = E[C_i] = \frac{1}{J} \sum_{i=1}^{L} C_i. \quad (28)$$

Regarding as (24), it is known that

$$R_c \cdot \log_2 M \cdot J \leq C = f(E_b/N_0). \quad (29)$$

When $R_c \cdot \log_2 M \cdot J = C$, the utilization of communication resources is maximized. The Shannon limit is defined by the minimum $E_b/N_0$ that can realize reliable transmission with data rate $R_c$, thus

$$(E_b/N_0)_{\text{min}} = f^{-1}(R_c \cdot \log_2 M \cdot J). \quad (30)$$

Nevertheless, the integral of (27) is extremely complex, especially its inverse function. Since it is one-to-one mapping between the code rate $R_c$ and $(E_b/N_0)_{\text{min}}$, the corresponding code rate $R_c$ can be calculated for a given $(E_b/N_0)_{\text{min}}$. Then, we can obtain a sequence of data points $(R_c, (E_b/N_0)_{\text{min}})$. 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 $(E_b/N_0)_{\text{min}}$ 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 $(E_b/N_0)_{\text{min}}$, because of the reduced minimum Euclidean distance of the superimposed signals.

**VI. SIMULATION RESULTS**

In this section, we simulate the AUER and BER performances of the proposed system. The UDAS sets are generated based on the cyclic mode, with generators $a_{L,4} = (1, 1i, 2, 2i)$ of $L = 4$, and $a_{L,6} = (1, 1i, 2, 2i, 4, 4i)$ of $L = 6$. Because of the cyclic structure, it is easy to know $T = L$. For convenience, we directly use $L = 4$ and $L = 6$ to stand for the UDAS sets generated by $a_{L,4}$ and $a_{L,6}$, respectively. Moreover, $G_1$ is a $(3,6)$-regular QC-LDPC (1016, 508) with rate 0.5, and $G_2$ is a SPC with rate $R_2 = 1 - \frac{1}{L \cdot \log_2 M}$. Thus, the total data rate of each user is equal to $R_e = 0.5 \times (1 - \frac{1}{L \cdot \log_2 M})$. For example, when $L = 4$ and $M = 2$, it is found that $R_e = 0.375$.

First of all, we focus on the AUER performance, with the assumption that the number of arrival users $J$ (or $\tau$) 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, where the UDAS is generated based on the cyclic mode with $a_{L,4}$.

![AUER of the proposed system with various parameters](image)

Fig. 4. AUER of the proposed system with various parameters, where $M_1 = 1$, $L = 4$, and $M = 20, 60, 102, 138$ and 196.

From Fig. 4, it is found 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 perfectly accord with our deduced theoretical results. When we set $E_b/N_0 = 0$ dB and 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 an AUER of $10^{-5}$, at this moment, the required $E_b/N_0$ is only 0 dB. This result is appealing for a random access network.

Nevertheless, when it is an extremely short packet transmission scenario, e.g., the transmit packet is smaller than 100, the proposed SoF-AUD may exist an error floor. At this moment, we can combine the proposed SoF-AUD with the classical MAP detection to improve the AUER.

The BER performance of the proposed system is shown in Fig. 5, where the length of the encoded codeword is 1016. Therefore, it is reasonable to assume that the number of arrival users and UDAS set have been perfectly detected.

![BER performance of the proposed system](image)

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 $\tau = 4$. 

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Nevertheless, the BER differences among maximum iteration number is 10. When BER is much better than the case of L = 4, since the number of arrival users of the second group is larger than that of the first group. In addition, when the number of arrival users (i.e., τ = 4) and the dimension (i.e., M = 4) are given, the BER of the case of L = 6 is worse than that of that of L = 4, according to the conclusion of Fig. 5 (a).

Fig. 6 shows the BER comparison among different systems. We compare our proposed system with the Walsh sequence, the SCMA codebook [19], and the 4 × 7 logical signature matrix (LSM) of UD ternary code [7]. Note that, the SEs of the Walsh sequence, SCMA codebook and LSM systems are respectively 1 and 1.5, and 1.75. Our proposed UDAS based system selects four cases for comparison. The case one is to set the parameters as L = T = 4, and M = 1 without encoder, indicating that the SE of the case one is 4. Subsequently, the case two inserts encoder into the case one, and the SE of the case two becomes as 1.5, because of the inserted channel code. At this moment, our proposed UDAS system selects four cases for comparison. The case one is to set the parameters as L = T = 4, and M = 1 without encoder, whose SE is 6. The case four inserts encoder into the case three, so that the SE of the case four becomes as 2.5, due to λ = T · Rc = 6 × (1/2 × 1/2) = 1.5. Similarly, the parameters of the case three are set to be L = T = 6, and M = 1 without encoder, whose SE is 6. The case four inserts encoder into the case three, so that the SE of the case four becomes as 2.5, due to λ = T · R = 6 × (1/2 × 1/2) = 1.5.

It is found that, the Walsh sequence provides the best BER performance in an AWGN multiple-access channel, whose BER performance is the same as the classical BPSK modulation. Nevertheless, our proposed cases (i.e., cases one and three) can provide higher SEs than the Walsh sequence, SCMA codebook and LSM sequence. Particularly, our proposed case one with λ = 4 has better BER performance than that of the LSM sequence with λ = 1.75. When E_b/N_0 is larger than 6.6 dB, our proposed case two can provide better BER performance than the Walsh sequence and SCMA codebook, because of the inserted channel code. At this moment, our proposed case two has the same SE as that of SCMA system, verifying the validity of our proposed system.
VII. CONCLUSION

This paper introduces UDAS for GFMA systems. We give some definitions of UDAS set, and construct two kinds of UDAS sets. Essentially, the proposed cyclic/quasi-cyclic structure can help the receiver realize low-complexity AUD. Then, we present an UDAS-based MD-BICM transceiver, consisting of a SoF-AUD and an iteration MUD algorithm. Both the theoretical AUER and Shannon limits are deduced in details. Simulation results show that, when $E_b/N_0 = 0$ dB, our proposed system can support four simultaneously arrival users with an extremely low AUER of $10^{-5}$, which is appealing for a random access network.

In the future, the proposed UDAS set can be applied to a general random access scenario, e.g., uncoordinated multiple-access. Additionally, there are many approaches to construct UDAS sets.

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 Commun. Mag., vol. 53, no. 9, pp. 74–81, Sep. 2015.

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

[3] J. Dai, K. Niu, Z. Si, C. Dong, and J. Lin, “Polar-coded non-orthogonal multiple access,” IEEE Trans. Signal Process., vol. 66, no. 5, pp. 1374–1389, May 2018.

[4] Y. Wu, X. Gao, S. Zhou, W. Yang, Y. Polyanskiy, and G. Caire, “Massive access for future wireless communication systems,” IEEE Wireless Commun., vol. 27, no. 4, pp. 148–156, Aug. 2020.

[5] J. Belfiore, G. Rekaya, and E. Viterbo, “The golden code: A $2 \times 2$ full rate space-time code with nonvanishing determinants,” IEEE Trans. Inf. 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,” in Proc. IEEE Global Commun. Conf., Dec. 2017, pp. 1–7.

[7] O. Mashayekhi and F. Marvasti, “Uniquely decodable codes with fast decoder for overloaded synchronous CDMA systems,” IEEE Trans. Commun., 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,” in Proc. Annu. IEEE India Conf. (INDICON), Dec. 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,” in Proc. IEEE Wireless Commun. Netw. Conf. (WCNC), Apr. 2019, pp. 1–7.

[10] M. Kulhandjian and D. A. Pado, “Uniquely decodable code-division via augmented Sylvester-Hadamard matrices,” in Proc. IEEE Wireless Commun. Netw. Conf. (WCNC), Apr. 2012, pp. 359–363.

[11] M. Li and Q. Liu, “Fast code design for overloaded code-division multiplexing systems,” IEEE Trans. Veh. Technol., 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,” in Proc. IEEE 29th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC), Sep. 2018, pp. 1–6.

[13] D. V. Sarwate and M. B. Pursley, “Cross correlation properties of pseudo random and related sequences,” Proc. IEEE, vol. 66, no. 5, pp. 593–619, May 1980.

[14] B. M. Popovic, “Generalized chirp-like polyphase sequences with optimum correlation properties,” IEEE Trans. Inf. Theory, vol. 38, no. 4, pp. 1406–1409, Jul. 1992.

[15] M. Hua et al., “Analysis of the frequency offset effect on random access signals,” IEEE Trans. Commun., vol. 61, no. 11, pp. 4728–4740, Nov. 2013.

[16] Y.-R. Tsai, H.-Y. Huang, Y.-C. Chen, and K.-J. Yang, “Simultaneous multiple carrier frequency offsets estimation for coordinated multi-point transmission in OFDM systems,” IEEE Trans. Wireless Commun., vol. 12, no. 9, pp. 4558–4568, Sep. 2013.

[17] R. Abbas, M. Shirvani Moghaddam, Y. Li, and B. Vucetic, “A novel analytical framework for massive grant-free NOMA,” IEEE Trans. Commun., vol. 67, no. 3, pp. 2436–2449, Mar. 2019.

[18] A. Bayesteh, E. Yi, H. Nikopour, and H. Baligh, “Blind detection of SCMA for uplink grant-free multiple-access,” in Proc. 11th Int. Symp. Wireless Commun. Syst. (ISWCS), Aug. 2014, pp. 853–857.

[19] Y. Zhou, Q. Yu, W. Meng, and C. Li, “SCMA codebook design based on constellation rotation,” in Proc. IEEE Int. Conf. Commun. (ICC), May 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,” in Proc. IEEE/CIC Int. Conf. Commun. China, Aug. 2018, pp. 112–116.

[21] X. Zhang, G. Han, D. Zhang, and L. Yang, “A lattice-based SCMA codebook design for IoMT communications,” in Proc. IEEE/CIC Int. Conf. Commun. Workshops China, Aug. 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 Commun. Lett., vol. 21, no. 9, pp. 1925–1928, Sep. 2017.

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

[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 Commun. Lett., 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 J. Sel. Areas Commun., 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 Process. Mag., vol. 35, no. 5, pp. 88–99, Sep. 2018.

[27] Q.-Y. Yu, W.-X. Meng, and S. Lin, “Packet loss recovery scheme with uniquely-decodable codes for streaming multimedia over P2P networks,” IEEE J. Sel. Areas Commun., vol. 31, no. 9, pp. 142–154, Sep. 2013.

[28] Q.-Y. Yu, H.-H. Chen, and W.-X. Meng, “A unified multiuser coding framework for multiuser uplink massive MIMO,” IEEE Syst. J., 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, U.K.: Cambridge Univ Press, 2017.

[30] T. M. Cover and J. A. Thomas, Elements of Information Theory, Beijing, China: Tsinghua Univ. Press, 2002.