Joint Active User Detection and Channel Estimation in Massive Access Systems Exploiting Reed-Muller Sequences

Jue Wang, *Student Member, IEEE,* Zhaoyang Zhang, *Member, IEEE,* and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—The requirements to support massive connectivity and low latency in massive Machine Type Communications (mMTC) bring a huge challenge in the design of its random access (RA) procedure, which usually calls for efficient joint active user detection and channel estimation. In this paper, we exploit the vast sequence space and the beneficial nested structure of the length-2^m second-order Reed-Muller (RM) sequences for designing an efficient RA scheme, which is capable of reliably detecting multiple active users from the set of unknown potential users with a size as large as $2^{m(m-1)/2}$, whilst simultaneously estimating their channel state information as well. Explicitly, at the transmitter each user is mapped to a specially designed RM sequence, which facilitates reliable joint sequence detection and channel estimation based on a single transmission event. To elaborate, as a first step, at the receiver we exploit the elegant nested structure of the RM sequences using a layer-by-layer RM detection algorithm for the single-user (single-sequence) scenario. Then an iterative RM detection and channel estimation algorithm is conceived for the multi-user (multi-sequence) scenario. As a benefit of the information exchange between the RM sequence detector and channel estimator, a compelling performance vs. complexity trade-off is struck, as evidenced both by our analytical and numerical results.**

***Index Terms*—mMTC, Massive access, Reed-Muller sequences, User identification, Channel estimation.**

I. Introduction

A. Motivation

The 5G wireless mobile network aims to support diverse applications, which were classified by 3GPP into three categories: enhanced Mobile Broadband (eMBB), Ultra-Reliable and Low Latency Communications (URLLC), and massive Machine Type Communications (mMTC) [1]. For the mMTC scenario, the key requirement is to provide massive connectivity for instant access of short and sporadic data traffic in applications like smart city and industrial internet of things [2].

The random access (RA) procedure is key to enable the massive connectivity. As depicted in Fig. 1 (a), in the current long term evolution (LTE) system, a four-message handshake between the user equipment (UE) and the access point (AP), which consists of Preamble, Random Access Response (RAR), Initial Layer 3 Message (Msg3) and Contention Resolution (CR) [3], is used for RA. If the UE successfully accesses the system, it completes the RA procedure by sending the "RRC Connection Setup Complete" message and starts to transmit its data. When applied to mMTC, this RA scheme faces two challenges. Firstly, the small-sized sequence space suffers from a high collision rate. When a UE triggers the RA procedure, it randomly chooses a preamble from the available sequence pool to send in the next RA slot. According to LTE [4], there are 54 Zadoff-Chu (ZC) sequences reserved for contention in each cell. If more than one UE select the same sequence, a collision occurs and these UEs must retransmit the sequence. In the context of massive connectivity, the frequent collisions and precipitated retransmissions will lead to network congestion, increased delay and resource wastage. Secondly, the traffic pattern exchanged in machine-type communications is typically short and sporadic [5]. A large amount of signalling will lead to high signalling-to-data ratio and low bandwidth efficiency. Since the ZC sequences do not carry any information about user identifiers (IDs) in the LTE system, the AP has no knowledge of exactly who is requiring access until Msg3 has been successfully received.

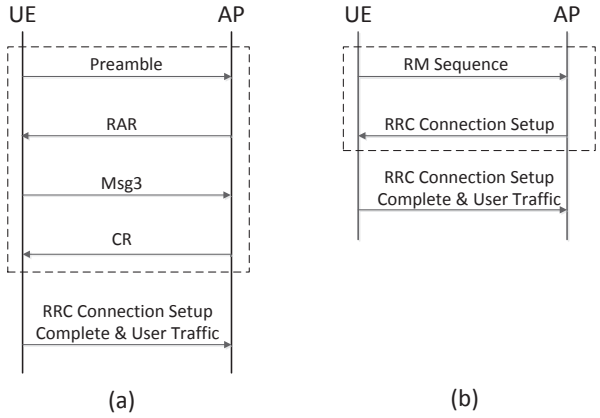


Fig. 1. (a) The RA procedure of LTE system. (b) The RA procedure using RM sequences.

Hence, the preamble sequences play a vital role in the RA procedure. To satisfy the tight specifications of mMTC, the sequence space has to match the number of users in the system. Furthermore, predefined sequences with user IDs embedded in them are preferred for the sake of low signalling overhead and

This work was supported in part by National Natural Science Foundation of China under Grant 61725104 and 61631003, Huawei Technologies Co., Ltd under Grant HF2017010003, YB2015040053 and YB2013120029, and the Fundamental Research Funds for the Central Universities.

Jue Wang (e-mail: juew@zju.edu.cn) and Zhaoyang Zhang (Corresponding Author, e-mail: zhzy@zju.edu.cn) are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China. Lajos Hanzo (e-mail: lh@ecs.soton.ac.uk) is with the Department of Electronics and Computer Science, University of Southampton, UK.

system latency. In this case, the AP has to detect the IDs of the active users and estimate the channel coefficients based on the received signal. To reduce congestion and achieve low latency, the above procedures have to be accomplished with high accuracy and at low complexity.

Moreover, it is worth pointing out that the difficulty of detecting active users from sets of potential users with different set sizes is very different. In general, detecting active users from a set of tens or hundreds of potential users can be readily achieved by elaborate sequence design and exhaustive search. However, it is rather challenging to do this in the massive access scenario where the number of unknown potential users can be as huge as hundreds of thousands or even millions, which makes the detection reminiscent of looking for the needle in a haystack.

B. Related Work

Numerous researchers have improved the RA procedures by designing beneficial preamble sequences. Some of them focus on increasing the number of available ZC sequences. For example, the authors of [6] proposed a preamble reuse scheme by partitioning the cell coverage and reducing the cyclic shift size. Apart from the classic ZC sequences, [7] they inserted extra auxiliary sequences in the positions randomly chosen from the first data sub-frame. However, the use of auxiliary sequences decreases the data transmission efficiency and increases the detection complexity. By contrast, the RA procedure of [8] treats several RA slots as a virtual frame and the active users randomly choose a ZC sequence for their transmission in each slot. This is equivalent to expanding the sequence space at the cost of an increased overhead ratio and detection complexity.

Other authors investigate different sequence generation methods in support of massive connectivity and active user detection. On one hand, considering the huge number of potential users in the system, it is impossible to assign orthogonal sequences to each user. On the other hand, non-orthogonal sequences would inevitably impose multi-user interference during the detection. Hence, the key challenge is to obtain a relatively large sequence space, whilst simultaneously attaining reliable detection. Based on the sparsity of user-activity caused by the sporadic traffic in mMTC, the active user detection is often formulated as a compressive sensing (CS) problem. The active users transmit a unique sequence to access the system. To satisfy the so-called Restricted Isometry Property (RIP) of CS [9], the entries of these sequences are usually generated either by Gaussian or Bernoulli processes [10]. Hence, considerable storage resources are required for storing these sequences at the AP. At the receiver, CS algorithms, such as Basis Pursuit (BP) [11], Orthogonal Matching Pursuit (OMP) [12] and approximate message passing (AMP) [13] [14], are utilized for detecting the active users. However, all these algorithms impose a high complexity that is related either to the size of sequence space or to the number of potential users in the system. Based on the code-expanded scheme of [8], the sequences constructed by Bloom filtering are utilized in [15]. Despite the lower latency compared to LTE, the hash functions used during the sequence generation have to be communicated both to all the users and to the AP, which imposes extra signalling overhead.

In [17], second-order RM sequences are invoked for massive access in 5G. Compared to ZC sequences, RM sequences have the following distinct advantages:

- The size of RM sequence space is several orders of magnitudes larger than that of the same-length ZC sequences, which meets the requirement of massive connectivity.
- The RM sequences can be uniquely and unambiguously mapped to user IDs. Hence, the AP can immediately infer the user ID right after the RM sequences have been correctly detected. Thus, the four-message handshake depicted in Fig. 1(a) is simplified to the twin-step interaction shown in Fig. 1(b), which results both in reduced signalling overhead and in reduced system latency.
- The structural properties of RM sequences can be exploited to design efficient detection algorithms having a low computational complexity.

In [16], RM sequences are used to construct the deterministic measurement matrix of CS, where an efficient reconstruction algorithm is also conceived, whose complexity is determined by the length of RM sequences and by the sparsity factor instead of the sequence space cardinality. However, this algorithm suffers from a serious performance degradation when the sparsity factor increases. Based on the algorithm in [16], a beneficial shuffle operation is advocated in [17] for improving the performance of RM detection, albeit at a high computational complexity. However, the compelling structural properties of RM sequences have not been fully exploited. In addition, to cope with the multi-sequence-induced interference, traditional successive interference cancellation (SIC) is adopted both in [16] and in [17], where the strongest signal is detected first and then subtracted from the received signal before detecting the next signal, regardless of whether the detected sequence is correct or not. Hence, once the sequence is incorrectly detected, it will affect the performance of all subsequent detection stages.

C. Main Contributions

Considering the above attractive merits of RM sequences, we exploit them to facilitate active user detection and channel estimation in massive access systems in this paper. At the transmitter, each active user maps its ID to a unique RM sequence and then sends it to the AP. Here, a novel mapping rule is proposed. On one hand, it helps to improve the sequence detection performance. On the other hand, it enables us to accomplish sequence detection and channel estimation at the same time, without invoking an extra channel estimation step as the algorithms in [16] and [17].

At the receiver, we firstly design a single-sequence detection algorithm based on the nested structure of RM sequences, which reveals the relationships between its sub-sequences. Exploiting this structural property in the sequence detection algorithm, we intend to improve the detection performance at a reduced complexity. Moreover, to cope with the potential

error propagation problem of the above algorithm, an enhanced algorithm based on multiple lists is also conceived. As for the case where multiple active users coexist in the system, an iterative RM detection and channel estimation algorithm is adopted. At each stage of an iteration, we focus on a particular user. Considering other signals as noise, the single-sequence detection algorithm is invoked for updating this user's sequence detection result. After all the sequences are updated, the results are fed into the channel estimator to improve the accuracy of channel estimates. Then the enhanced channel estimates are utilized in the next iteration to reinforce the sequence detection.

Against the above background, our main contributions are:

- A novel mapping rule between user IDs and RM sequences is proposed, which helps to improve the sequence detection capability and simplify our high-accuracy channel estimation.
- We derive an explicit relationship between the RM sequence and its sub-sequences, which is referred to as *nested structure*. We exploit this structural property for conceiving a low-complexity algorithm for single sequence detection, which exhibits superior detection performance and channel estimation accuracy. In order to mitigate the error propagation, a moderate-complexity enhanced algorithm is also conceived.
- Iterative active-user detection and channel estimation scheme is then also proposed for the multi-sequence scenario. On one hand, the iterative detection may reduce the impact of incorrectly detected sequences. On the other hand, owing to the bi-directional information exchange between the active-user detector and channel estimator, the detection capability and channel estimation accuracy can mutually reinforce each other.

D. Paper Organization and Notations

The rest of the paper is organized as follows. Section II describes the system model and introduces the transmitter and receiver structure of RA scheme using RM sequences. The nested structure of second-order RM sequences is derived in Section III. The blind active user detection and channel estimation algorithms are given in Section IV. Section V depicts the analytical model of the proposed algorithms and outlines the properties of them. Simulation Results are given in Section VI. And finally, Section VII concludes the paper.

Throughout this paper, scalars are represented in lowercase letters. Boldface lowercase and uppercase letters denote column vectors and matrices, respectively. $\mathbf{0}^N$ and $\mathbf{1}^N$ are $N \times 1$ column vectors of all zeros and all ones, respectively. We use $(\cdot)^T$ to denote transpose of a matrix or vector and $(\cdot)^*$ to denote complex conjugate. $\|\cdot\|$ is the Euclidean norm. The symbol $\odot$ represents element-wise product and $\oplus$ represents element-wise modulo-2 addition. $x \sim \mathcal{CN}(\mu, \sigma^2)$ indicates that x is a complex Gaussian random variable with mean μ and variance σ^2. $\mathcal{O}(\cdot)$ is reserved for complexity estimation.

II. System Model

The massive uplink connectivity scenario illustrated in Fig. 2 is considered, where a single AP is serving a huge pool of C users and the maximum value of C depends on the sequence space size. All the users are assumed to be synchronized and each one of them has a unique user ID. In a specific RA slot, there are K active users in the system and the active user set is denoted as $\kappa \triangleq \{1, 2, \ldots, K\}$. Due to the fact that the traffic pattern in mMTC scenario is typically sporadic, it is satisfied that $K \ll C$.

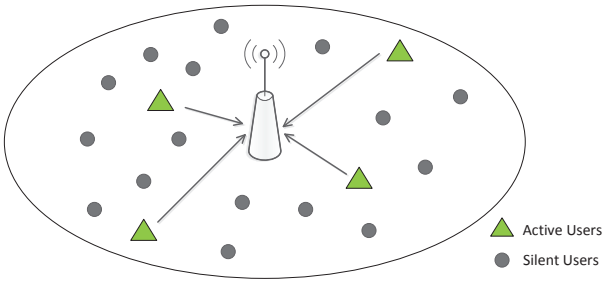


Fig. 2. The system model of the uplink massive connectivity scenario, where only a subset of users are active because of the sporadic traffic in mMTC.

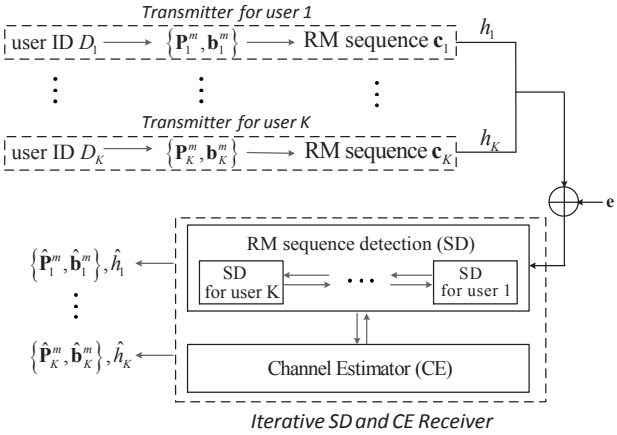


Fig. 3. The transmitter and receiver structure of random access scheme using RM sequences.

The transmitter and receiver structure is shown in Fig. 3. To access the system, every active user will simultaneously send an RM sequence along with its ID to the AP in the next available RA slot. The AP has to accomplish both active user detection and channel estimation based on the received signal. In the sequel, the processes at the transmitter and the receiver are described in detail, respectively.

A length-2^m second-order RM sequence $\mathbf{c}^m$ is determined by an $(m \times m)$-element symmetric binary matrix $\mathbf{P}^m$ and an $(m \times 1)$-element binary vector $\mathbf{b}^m$. According to [16], given the matrix-vector pair $\{\mathbf{P}^m, \mathbf{b}^m\}$, the j-th entry of the RM sequence $\mathbf{c}^m$ is expressed as

$$c_j^m = \frac{(-1)^{\mathrm{wt}(\mathbf{b}^m)}}{\sqrt{2^m}} i^{(2\mathbf{b}^m + \mathbf{P}^m \mathbf{a}_{j-1}^m)^T \mathbf{a}_{j-1}^m}, j = 1, \ldots, 2^m, \quad (1)$$

where i is the imaginary unit that satisfies $i^2 = -1$, $\mathrm{wt}(\mathbf{b}^m)$ denotes the Hamming weight of the vector $\mathbf{b}^m$, i.e., the number of 1s in $\mathbf{b}^m$, and $\mathbf{a}_{j-1}^m$ is the m-bit binary expression of $(j-1)$.

To generate its RM sequence, the active user $k, k\in\kappa$, first maps the user ID $D_k \in \{0, 1, \ldots, C-1\}$ to the matrix-vector pair $\{\mathbf{P}_k^m, \mathbf{b}_k^m\}$. We propose a novel mapping rule as follows:

1) Convert the user ID D_k to the length-$\lceil m(m-1)/2 \rceil$ binary vector $\left[d_{k,1}, \ldots, d_{k,m(m-1)/2}\right]$.
2) Put these bits successively into the symmetric matrix $\mathbf{P}_k^m$ having zero diagonal elements, i.e.,

$$\mathbf{P}_k^m = \begin{bmatrix} 0 & p_{k,(m,1)} & \cdots & p_{k,(m,m-1)} \\ p_{k,(m,1)} & 0 & \cdots & p_{k,(m-1,m-2)} \\ \vdots & \vdots & \ddots & \vdots \\ p_{k,(m,m-1)} & p_{k,(m-1,m-2)} & \cdots & 0 \end{bmatrix} = \begin{bmatrix} 0 & d_{k,1} & \cdots & d_{k,m-1} \\ d_{k,1} & 0 & \cdots & d_{k,2m-3} \\ d_{k,2} & d_{k,m} & \cdots & d_{k,3m-6} \\ \vdots & \vdots & \ddots & \vdots \\ d_{k,m-1} & d_{k,2m-3} & \cdots & 0 \end{bmatrix}.$$

3) The vector $\mathbf{b}_k^m = [b_{k,m}, b_{k,m-1}, \cdots, b_{k,1}]^T$ is constructed according to

$$b_{k,s} = \begin{cases} p_{k,(s,1)} \oplus p_{k,(s,2)} \oplus \ldots \oplus p_{k,(s,s-1)}, & 2 \leqslant s \leqslant m \\ b_{k,m} \oplus b_{k,m-1} \oplus \cdots \oplus b_{k,2}, & s = 1 \end{cases}. \tag{2}$$

Finally, the active user k generates the RM sequence $\mathbf{c}_k^m$ with the mapped matrix-vector pair $\{\mathbf{P}_k^m, \mathbf{b}_k^m\}$. According to the above mapping rule, both the Hamming weight of the vector $\mathbf{b}_k^m$ and the term $(\mathbf{P}_k^m \mathbf{a}_{j-1}^m)^T \mathbf{a}_{j-1}^m$ must be even. With the normalization coefficient $1/\sqrt{2^m}$ being omitted, the generation function shown in (1) can be simplified to

$$c_{k,j}^m = (-1)^{(\mathbf{b}_k^m)^T \mathbf{a}_{j-1}^m + \frac{1}{2}(\mathbf{a}_{j-1}^m)^T \mathbf{P}_k^m \mathbf{a}_{j-1}^m}. \tag{3}$$

The size of the sequence space created by the proposed mapping rule is up to $2^{m(m-1)/2}$ and the same-sized user space can be supported. In this sense, RM sequences are more suitable for massive access in mMTC scenario. Furthermore, the above mapping method also brings benefits for channel estimation and sequence detection, which will be specified in Section V.

All the active users in this RA slot simultaneously send their RM sequences to the AP and the received signal can be expressed as

$$y_j^m = \sum_{k=1}^{K} h_k \cdot c_{k,j}^m + e_j^m, \ j = 1, \ldots, 2^m, \tag{4}$$

where h_k represents the channel coefficient between the active user k and the AP, which is modeled as a complex Gaussian random process with unit variance, and e_j is the complex additive white Gaussian noise (AWGN) having mean zero and variance N_0. After receiving the signal, the AP has to detect all K active users and estimate the corresponding channel coefficients. Thanks to the one-to-one mapping between the matrix-vector pair and the user ID, active user detection can be accomplished by recovering the matrix-vector pairs $\{\mathbf{P}_k^m, \mathbf{b}_k^m\}$ from the received signal.

III. The nested structure of RM sequences

In this section, we introduce an important structural property of RM sequences, which is referred to as *nested structure* and is the fundamental of the following proposed RM detection algorithms. For ease of exposition, several relevant definitions are given as follows.

Definition 1: Given the matrix-vector pair $\{\mathbf{P}^m, \mathbf{b}^m\}$ and the RM sequence $\mathbf{c}^m$ generated by them, the lower right $(s \times s)$-element sub-matrix of $\mathbf{P}^m$ is defined as its *s-th order sub-matrix* $\mathbf{P}^s$ and the vector $\boldsymbol{\alpha}^s \triangleq [p_{s,1}, \ldots, p_{s,s-1}]^T$ denotes the *s-th layer* of $\mathbf{P}^m$. Furthermore, the *s-th order sub-vector* of $\mathbf{b}^m$ is defined as $\mathbf{b}^s \triangleq [b_s, b_{s-1}, \ldots, b_1]^T$. Then, the matrix-vector pair can be expressed in the following recursive form

$$\mathbf{P}^s = \begin{bmatrix} 0 & (\boldsymbol{\alpha}^s)^T \\ \boldsymbol{\alpha}^s & \mathbf{P}^{s-1} \end{bmatrix}, \quad \mathbf{b}^s = \begin{bmatrix} b_s & \mathbf{b}^{s-1} \end{bmatrix}. \tag{5}$$

The resultant length-2^s RM sequence determined by $\{\mathbf{P}^s, \mathbf{b}^s\}$ according to (3) is then termed as the *s-th order sub-sequence* $\mathbf{c}^s$.

The nested structure of RM sequences is summarized in the following theorem.

Theorem 1: As depicted in Fig. 4, given the RM sequence $\mathbf{c}^m$, its sub-sequences of order s and $(s-1)$ satisfy the *nested structure* that

$$\mathbf{c}^s = (\mathbf{c}^{s-1}, \mathbf{c}^{s-1} \odot \mathbf{v}^{s-1}), 2 \leq s \leq m. \tag{6}$$

The vector $\mathbf{v}^{s-1}$ is a length-2^{s-1} Walsh sequence determined by the s-th layer of the matrix $\mathbf{P}^m$, and its j-th entry is given by

$$v_j^{s-1} = (-1)^{b_s} \cdot (-1)^{(\boldsymbol{\alpha}^s)^T \mathbf{a}_{j-1}^{s-1}} = (-1)^{(\boldsymbol{\alpha}^s)^T \overline{\mathbf{a}_{j-1}^{s-1}}}, \tag{7}$$

where $\overline{\mathbf{a}_{j-1}^{s-1}} \triangleq \mathbf{a}_{j-1}^{s-1} \oplus \mathbf{1}^{s-1}$.

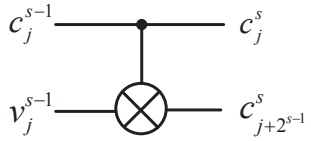


Fig. 4. The nested structure of RM sequences.

Proof: We prove that RM sequences obey the above nested structure relying on the generation function given in (3). When we have $1 \leq j \leq 2^{s-1}$, the vector $\mathbf{a}_{j-1}^s$ can be decomposed as

$$\mathbf{a}_{j-1}^s = \begin{bmatrix} 0 & \mathbf{a}_{j-1}^{s-1} \end{bmatrix}^T.$$

Then the exponent of c_j^s is expanded as

$$\begin{aligned} &(\mathbf{b}^s)^T \mathbf{a}_{j-1}^s + \frac{1}{2}(\mathbf{a}_{j-1}^s)^T \mathbf{P}^s \mathbf{a}_{j-1}^s \\ &= \begin{bmatrix} b_s & (\mathbf{b}^{s-1})^T \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{a}_{j-1}^{s-1} \end{bmatrix} \\ &\quad + \frac{1}{2} \begin{bmatrix} 0 & (\mathbf{a}_{j-1}^{s-1})^T \end{bmatrix} \begin{bmatrix} 0 & (\boldsymbol{\alpha}^s)^T \\ \boldsymbol{\alpha}^s & \mathbf{P}^{s-1} \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{a}_{j-1}^{s-1} \end{bmatrix} \end{aligned}$$

$$=(\mathbf{b}^{s-1})^T\mathbf{a}_{j-1}^{s-1}+\frac{1}{2}(\mathbf{a}_{j-1}^{s-1})^T\mathbf{P}^{s-1}\mathbf{a}_{j-1}^{s-1},$$

hence we have $c_j^s = c_j^{s-1}$ in this case. Similarly, by decomposing $\mathbf{a}_{j+2^{s-1}-1}^s$ as

$$\mathbf{a}_{j+2^{s-1}-1}^s = \begin{bmatrix} 1 & \mathbf{a}_{j-1}^{s-1} \end{bmatrix}^T,$$

it is derived that

$$\begin{aligned}&(\mathbf{b}^s)^T\mathbf{a}_{j-1}^s+\frac{1}{2}(\mathbf{a}_{j-1}^s)^T\mathbf{P}^s\mathbf{a}_{j-1}^s\\ =&(\mathbf{b}^{s-1})^T\mathbf{a}_{j-1}^{s-1}+\frac{1}{2}(\mathbf{a}_{j-1}^{s-1})^T\mathbf{P}^{s-1}\mathbf{a}_{j-1}^{s-1}+b_s+(\boldsymbol{\alpha}^s)^T\mathbf{a}_{j-1}^{s-1}.\end{aligned}$$

Therefore we have $c_{j+2^{s-1}}^s = c_j^{s-1}\cdot v_j^{s-1}$, where

$$v_j^{s-1} = (-1)^{b_s}\cdot(-1)^{(\boldsymbol{\alpha}^s)^T\mathbf{a}_{j-1}^{s-1}}.$$

Combined with $b_s = (\boldsymbol{\alpha}^s)^T\mathbf{1}^{s-1}$ from (2), it is recast that $v_j^{s-1} = (-1)^{(\boldsymbol{\alpha}^s)^T\overline{\mathbf{a}_{j-1}^{s-1}}}$, where $\overline{\mathbf{a}_{j-1}^{s-1}} \triangleq \mathbf{a}_{j-1}^{s-1}\oplus\mathbf{1}^{s-1}$.

On this basis, we may conclude that

$$\begin{cases} c_j^s = c_j^{s-1} \\ c_{j+2^{s-1}}^s = c_j^{s-1}\cdot v_j^{s-1} \end{cases}, 1\le j\le 2^{s-1}, \tag{8}$$

which completes the proof. ∎

IV. Blind Active User Detection and Channel Estimation Algorithm

In this section, a novel RM detection algorithm is proposed for single sequence detection. Based on the nested structure of RM sequences, we recover the matrix-vector pair layer by layer and estimate the channel coefficients in a straightforward manner. Moreover, an enhanced algorithm is conceived for handling error propagation. In the case where multiple RM sequences coexist, an iterative RM detection and channel estimation algorithm is adopted.

A. Single-Sequence Scenario: Layer-by-Layer RM Detection Algorithm

When there is only a single active user in the system, the received signal can be rewritten as

$$y_j^m = h\cdot c_j^m + e_j^m, j = 1,\dots,2^m.$$

The layer-by-layer RM detection algorithm is executed following the process shown in Fig. 5.

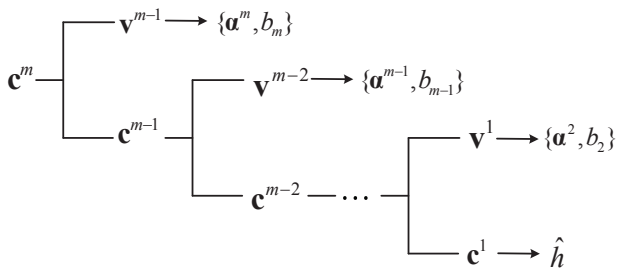


Fig. 5. The process of layer-by-layer RM detection algorithm.

Firstly, we split the received signal into two length-2^{m-1} parts and denote them as $(\mathbf{y}^m)' = \left[y_1^m,\dots,y_{2^{m-1}}^m\right]^T$ and $(\mathbf{y}^m)'' = \left[y_{2^{m-1}+1}^m,\dots,y_{2^m}^m\right]^T$, respectively. According to the nested structure summarized in (8), the elements of these two partial sequences are given by

$$\begin{cases} (y_j^m)' = hc_j^{m-1}+e_j^m \\ (y_j^m)'' = hc_j^{m-1}\cdot v_j^{m-1}+e_{j+2^{m-1}}^m \end{cases}, j = 1,\dots,2^{m-1}.$$

Then, we perform the element-wise conjugate multiplication on them, i.e.,

$$\tilde{y}_j^m = (y_j^m)'\cdot\left((y_j^m)''\right)^* = |h|^2\cdot v_j^{m-1}+\tilde{\varepsilon}_j^m, \tag{9}$$

where

$$\tilde{\varepsilon}_j^m \triangleq hc_j^{m-1}(e_{j+2^{m-1}}^m)^* + h^*c_j^{m-1}v_j^{m-1}e_j^m + e_j^m(e_{j+2^{m-1}}^m)^*.$$

As shown in (7), $\mathbf{v}^{m-1}$ is a Walsh sequence with frequency $\boldsymbol{\alpha}^m$. Hence, $\boldsymbol{\alpha}^m$ can be recovered by performing the Walsh-Hadamard Transformation (**WHT**) on (9). Specifically, the m-th order Hadamard matrix is represented as $\mathbf{T}^m = [\mathbf{t}_1^m,\dots,\mathbf{t}_i^m,\dots,\mathbf{t}_{2^m}^m]^T$ and its elements are $t_{i,j}^m = (-1)^{(\mathbf{a}_{i-1}^m)^T\overline{\mathbf{a}_{j-1}^m}}, i,j\in\{1,2,\dots,2^m\}$. Note that the Hadamard matrix we used here is obtained by flipping the normalized one, whose element is $(-1)^{(\mathbf{a}_{i-1}^m)^T\mathbf{a}_{j-1}^m}$, up and down. We denote the result of the transformation by the vector $\mathbf{V}^{m-1} = \mathbf{WHT}\{\tilde{\mathbf{y}}^m\} = \mathbf{T}^{m-1}\cdot\tilde{\mathbf{y}}^m$, whose i-th entry is

$$\begin{aligned} V_i^{m-1} =&\mathbf{t}_i^{m-1}\cdot\tilde{\mathbf{y}}^m \\ =&|h|^2\sum_{j=1}^{2^{m-1}}(-1)^{(\boldsymbol{\alpha}^m)^T\overline{\mathbf{a}_{j-1}^{m-1}}}\cdot(-1)^{(\mathbf{a}_{i-1}^{m-1})^T\overline{\mathbf{a}_{j-1}^{m-1}}}+\varepsilon_i^{m-1} \\ =&|h|^2\sum_{j=1}^{2^{m-1}}(-1)^{(\boldsymbol{\alpha}^m+\mathbf{a}_{i-1}^{m-1})^T\overline{\mathbf{a}_{j-1}^{m-1}}}+\varepsilon_i^{m-1}, \end{aligned} \tag{10}$$

where $\varepsilon_i^{m-1}\triangleq\sum_{j=1}^{2^{m-1}}(-1)^{(\mathbf{a}_{i-1}^{m-1})^T\overline{\mathbf{a}_{j-1}^{m-1}}}\cdot\tilde{\varepsilon}_j^m$. Thus, when we have $\mathbf{a}_{i-1}^{m-1} = \boldsymbol{\alpha}^m$, V_i^{m-1} will achieve its maximum value, which equals to $2^{m-1}|h|^2+\varepsilon_i^{m-1}$. On this basis, the vector $\boldsymbol{\alpha}^m$ can be recovered by searching through the vector $\mathbf{V}^{m-1}$ for the component having the largest value. Additionally, for simplicity, we can opt for comparing the real parts of $\mathbf{V}^{m-1}$ with no performance loss. Specifically, with

$$w = \arg\max_i\left\{(V_1^{m-1})_I,\dots,(V_i^{m-1})_I,\dots,(V_{2^{m-1}}^{m-1})_I\right\},$$

and $(\cdot)_I$ representing the real component, the estimate $\hat{\boldsymbol{\alpha}}^m$ is the $(m-1)$-bit binary expression of $(w-1)$ and then $\hat{b}^m$ is the modulo-2 addition of the entries in $\hat{\boldsymbol{\alpha}}^m$, i.e.,

$$\hat{\boldsymbol{\alpha}}^m = \mathbf{a}_{w-1}^{m-1},\quad \hat{b}_m = \hat{\alpha}_1^m\oplus\hat{\alpha}_2^m\oplus\cdots\oplus\hat{\alpha}_{m-1}^m. \tag{11}$$

Upon substituting (11) into (7), the estimate $\hat{\mathbf{v}}^{m-1}$ is obtained.

Next, combining the estimate $\hat{\mathbf{v}}^{m-1}$ and the pair of partial sequences $(\mathbf{y}^m)'$ and $(\mathbf{y}^m)''$, we arrive at:

$$\mathbf{y}^{m-1} = \frac{1}{2}\left((\mathbf{y}^m)'+\hat{\mathbf{v}}^{m-1}\odot(\mathbf{y}^m)''\right). \tag{12}$$

Under the condition that the estimate $\hat{\mathbf{v}}^{m-1}$ is correct, it is derived that

$$y_j^{m-1} = hc_j^{m-1}+e_j^{m-1}, j = 1,\dots,2^{m-1}, \tag{13}$$

where $e_j^{m-1} \triangleq \frac{1}{2}\left(e_j^m + \hat{v}_j^{m-1} \cdot e_{j+2^{m-1}}^m\right)$.

In the sequel, the sequence $\mathbf{y}^{m-1}$ is split into $\left(\mathbf{y}^{m-1}\right)'$ and $\left(\mathbf{y}^{m-1}\right)''$ and the estimates $\left\{\hat{\boldsymbol{\alpha}}^{m-1}, \hat{b}_{m-1}\right\}$ are calculated by repeating the above steps. This process continues until all the estimates $\left\{\hat{\boldsymbol{\alpha}}^s, \hat{b}_s\right\}, s \in \{2, \ldots, m\}$ and the length-2 sequence $\mathbf{y}^1 = \left[y_1^1, y_2^1\right]^T$ are obtained.

Then, the channel estimation can be done based on $\mathbf{y}^1$. Assuming that all the previous estimates are correct, we have

$$\begin{cases} y_1^1 = hc_1^1 + e_1^1 = h + e_1^1 \\ y_2^1 = hc_2^1 + e_2^1 = h \cdot (-1)^{b_1} + e_2^1 \end{cases}, \tag{14}$$

where

$$\begin{cases} e_1^1 = \frac{1}{2^{m-1}}\left(e_1^m \pm e_3^m \pm \cdots \pm e_{2^m-1}^m\right) \\ e_2^1 = \frac{1}{2^{m-1}}\left(e_2^m \pm e_4^m \pm \cdots \pm e_{2^m}^m\right) \end{cases},$$

and the signs depend on the estimates $\hat{\mathbf{v}}^s, 2 \le s \le m-1$. Since $e_j^m, j = 1, \ldots, 2^m$ are i.i.d. random variables following the distribution $\mathcal{CN}(0, N_0)$, the channel coefficient can thus be estimated by

$$\hat{h} = \frac{1}{2}\left(y_1^1 + (-1)^{\hat{b}_1} \cdot y_2^1\right), \tag{15}$$

where $\hat{b}_1 = \hat{b}_m \oplus \cdots \oplus \hat{b}_2$ according to (2).

Capitalizing on all the estimations above, the matrix-vector pair $\{\mathbf{P}^m, \mathbf{b}^m\}$ is recovered according to (5) and then reverts to the active user ID. To summarize, the layer-by-layer RM detection algorithm conceived for single sequence detection is presented in **Algorithm 1**.

Algorithm 1 Layer-by-Layer RM Detection Algorithm

Input: the received signal $\mathbf{y}^m$;
Output: the matrix-vector pair $\{\hat{\mathbf{P}}^m, \hat{\mathbf{b}}^m\}$, the channel coefficient $\hat{h}$.

1: **for** $s = m : -1 : 2$ **do**
2: Split $\mathbf{y}^s$ into two partial sequences $(\mathbf{y}^s)'$ and $(\mathbf{y}^s)''$.
3: Perform the element-wise conjugate multiplication on the $(\mathbf{y}^s)'$ and $(\mathbf{y}^s)''$ with (9).
4: Perform the **WHT** and obtain the estimates $\{\hat{\boldsymbol{\alpha}}^s, \hat{b}_s\}$ according to (11).
5: Obtain $\hat{\mathbf{v}}^{s-1}$ by applying $\{\hat{\boldsymbol{\alpha}}^s, \hat{b}_s\}$ to (7).
6: Calculate the sequence $\mathbf{y}^{s-1}$ with (12).
7: **end for**
8: With the element $\hat{b}_1 = \hat{b}_m \oplus \cdots \oplus \hat{b}_2$, the channel coefficient is estimated according to (15).
9: The matrix-vector pair $\{\hat{\mathbf{P}}^m, \hat{\mathbf{b}}^m\}$ is recovered based on all the estimates $\left\{\hat{\boldsymbol{\alpha}}^s, \hat{b}_s, \hat{b}_1\right\}, s \in \{2, \ldots, m\}$ and then reverts to the user ID.
10: **return** $\{\hat{\mathbf{P}}^m, \hat{\mathbf{b}}^m\}, \hat{h}$.

However, we found that Algorithm 1 suffers from an error propagation problem. To elaborate, let us assume that an error occurs during the recovery of $\{\boldsymbol{\alpha}^s, b_s\}$, which results in $\hat{\mathbf{v}}^{s-1} \neq \mathbf{v}^{s-1}$. Then the result of (12) turns out to be

$$\begin{aligned} y_j^{s-1} &= \frac{1}{2}\left(\left(y_j^s\right)' + \hat{v}_j^{s-1} \cdot \left(y_j^s\right)''\right) \\ &= \frac{1}{2} h c_j^{s-1}\left(1 + v_j^{s-1}\hat{v}_j^{s-1}\right) + e_j^{s-1} \\ &= \begin{cases} hc_j^{s-1} + e_j^{s-1}, & \text{if } v_j^{s-1} = \hat{v}_j^{s-1} \\ e_j^{s-1}, & \text{if } v_j^{s-1} \neq \hat{v}_j^{s-1} \end{cases}. \end{aligned} \tag{16}$$

Since $\hat{\mathbf{v}}^{s-1}$ and $\mathbf{v}^{s-1}$ are two different Walsh sequences, the cardinality of the set $J = \left\{j | v_j^{s-1} \neq \hat{v}_j^{s-1}\right\}$ definitely equals to 2^{s-2}. As a result, the sequence $\mathbf{y}^{s-1}$ loses half the information of $\mathbf{c}^{s-1}$, which will impose consistent errors in the following steps.

In the sequel, we propose a list detection algorithm for overcoming this error propagation problem. The algorithm can be partitioned into the extension and validation stages.

Firstly, let us consider the extension stage. For ease of clarification, we introduce two parameters $(\mathbf{L}, F)$, where $\mathbf{L} = [\, L_m, L_{m-1}, \ldots, L_{m-F+1}]$ determines the width of extension and $F = |\mathbf{L}|$, i.e., the number of elements in the vector $\mathbf{L}$, determines the depth of extension. On this basis, during the recovery of $\{\boldsymbol{\alpha}^s, b_s\}$ (line 4 of Algorithm 1), if $s \ge m - F + 1$, instead of choosing the location having the largest real component, we retain the L_s most likely locations to obtain a series of candidates. Then, the detection is continued in these parallel threads, each called a detection path. For the list detection algorithm with parameters $(\mathbf{L}, F)$, the total number of the detection paths at the end equals to $\prod_{s=m}^{m-F+1} L_s$, and the path having the notation $\mathbf{l} = (l_m, \ldots, l_{m-f+1}, \ldots, l_{m-F+1})$ means that we choose the l_{m-f+1}-th list in the f-th layer along this path. Intuitively, if $F = m-1$ and $\mathbf{L} = \left[\, 2^{m-1}, 2^{m-2}, \ldots, 2\right]$, the list detection algorithm is equivalent to the classic maximum likelihood sequence estimation (MLSE), which achieves the optimal performance with a complexity that is super-linear to the size of the sequence space. Faced with the huge sequence space of RM sequences, the complexity of MLSE is excessive. In the proposed list detection algorithm, we keep track of several most likely paths to reach a compromise between complexity and performance. An example of $([2, 2], 2)$ list detection algorithm is illustrated in Fig. 6.

At the validation stage, the optimal path is chosen by measuring the residual sequence energy. To be specific, the residual sequence energy on the path $\mathbf{l}$ is expressed as

$$R_{\mathbf{l}} = \left\| \mathbf{y}^m - \hat{h}_{\mathbf{l}} \cdot \hat{\mathbf{c}}_{\mathbf{l}}^m \right\|^2. \tag{17}$$

Then the path having the minimum residual energy is decided to be the final detection result. The reasons to use the residual energy as the path metric are two fold. On one hand, it is intuitive that if the detected sequence actually exists, the residual energy should be reduced after it is canceled from the received signal. On the other hand, if errors occur during the detection process, the channel coefficient estimated by (15) can be proved to be a very small value. Specifically, if the algorithm is executed to the last step based on the result

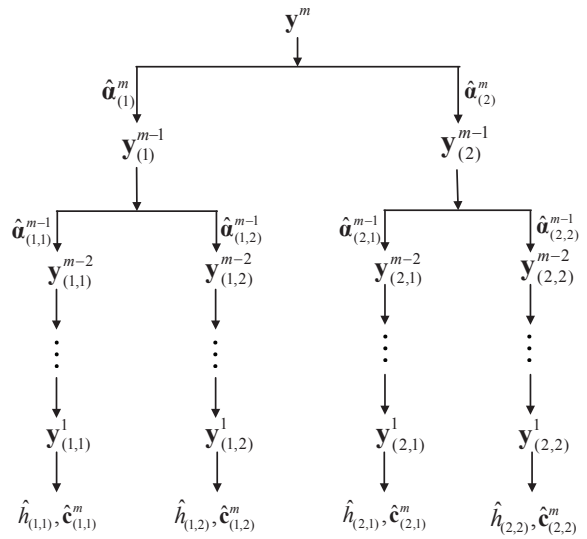


Fig. 6. The example of list detection algorithm with $\mathbf{L} = [2, 2]$ and $F = 2$.

derived in (16), the sequence $\mathbf{y}_1^1$ would have the form

$$y_j^1 = \begin{cases} hc_j^1 + e_j^1, & \text{if } v_j^{s'} = \hat{v}_j^{s'}, \forall s' = s-1, \ldots, 2 \\ e_j^1, & \text{otherwise} \end{cases}.$$

Since there is little chance that $v_j^{s'} = \hat{v}_j^{s'}$ is satisfied for every $s' \in \{s-1, \ldots, 2\}$ when errors occurred in the previous layers, we have $y_j^1 \approx e_j^1 \approx 0$ and then $\hat{h} \approx 0$ according to (15). Hence, the result calculated by (17) is almost equal to the energy of the received signal and the corresponding path will be discarded.

B. Multi-Sequence Scenario: Iterative RM Detection and Channel Estimation Algorithm

In this section, we consider the scenario where multiple active users coexist in the system. We detect the RM sequences and estimate the channel coefficients in an iterative manner. The process is described in detail below.

We first introduce some notations. The maximum number of detected users is limited to $k_{\max}$. On one hand, a small value of $k_{\max}$ would undoubtedly result in the unsatisfactory successful detection probability. On the other hand, if $k_{\max}$ is set to be too large, it may cause high false alarm rate (FAR) and computational complexity. Hence, the choice of the exact value of $k_{\max}$ is determined by the trade-off between the successful detection probability and the tolerable FAR, as well as the practical complexity constraint of the system. For ease of exposition, the superscripts in $\mathbf{y}^m$ and $\mathbf{c}^m$ are dropped. The stage where the RM sequence of the detected user k is updated during the n-th iteration is denoted as Stage (n, k). Furthermore, $\mathbf{c}_k^{(n)}$ and $h_k^{(n)}$ represent the RM detection result and the channel estimate at Stage (n, k), respectively.

Firstly, we initialize the RM sequence and the channel coefficient as $h_k^{(0)} = 0$ and $\mathbf{c}_k^{(0)} = \mathbf{0}^{2^m}$, respectively. At Stage (n, k), the multiple access interference (MAI) is expressed as

$$\zeta_{k,j}^{(n)} = \sum_{k' \neq k} h_{k'}^{(n')} \cdot c_{k',j}^{(n')}, \tag{18}$$

where

$$\begin{cases} n' = n, & \text{if } 1 \leq k' \leq k-1 \\ n' = n-1, & \text{if } k+1 \leq k' \leq k_{\max} \end{cases}.$$

Then, the result of $(y_j - \zeta_{k,j}^{(n)})$ is input to **Algorithm 1** to obtain $\mathbf{c}_k^{(n)}$ and $h_k^{(n)}$.

Next, after all the detected users' RM sequences are updated in this iteration, the results are fed to the linear least square channel estimator for further improving their accuracy, which is specified as

$$\begin{aligned} \mathbf{h}^{(n)} &= \left[h_1^{(n)}, \ldots, h_{k_{\max}}^{(n)}\right]^T \\ &= \underset{[h_1, \ldots, h_{k_{\max}}]^T \in \mathbb{C}^{k_{\max}}}{\arg\min} \sum_{j=1}^{2^m} \left\| y_j - \sum_{k'=1}^{k_{\max}} h_{k'} \cdot c_{k',j}^{(n')} \right\|^2. \end{aligned} \tag{19}$$

Thanks to the information exchange between the sequence detector and the channel estimator, the performance gradually improves and the algorithm is terminated when the results are converged or the maximum number of iteration is reached. The multi-sequence detection procedure is summarized in **Algorithm 2**.

Algorithm 2 Iterative RM Detection and Channel Estimation Algorithm

Input:
the received signal $\mathbf{y}$,
the maximum number of iteration $n_{\max}$,
the maximum number of detected users $k_{\max}$.
Output:
the active user ID set $\mathbf{D} = \{D_1, \ldots, D_{k_{\max}}\}$,
the channel coefficient $\hat{\mathbf{h}}$.

1: Initialization:
$\mathbf{C}^{(0)} = \left[\mathbf{c}_1^{(0)}, \ldots, \mathbf{c}_{k_{\max}}^{(0)}\right] = \mathbf{0}^{2^m \times k_{\max}}$, $\mathbf{h}^{(0)} = \mathbf{0}^{k_{\max}}$.
2: **for** $n = 1 : n_{\max}$ **do**
3: **for** $k = 1 : k_{\max}$ **do**
4: Calculate the MAI $\zeta_{k,j}^{(n)}$ according to (18).
5: Input the result of $(y_j - \zeta_{k,j}^{(n)})$ to **Algorithm 1** and record the output as $\mathbf{c}_k^{(n)}$ and $h_k^{(n)}$.
6: **end for**
7: Combine the RM detection results obtained in this iteration to the matrix $\mathbf{C}^{(n)} = \left[\mathbf{c}_1^{(n)}, \ldots, \mathbf{c}_{k_{\max}}^{(n)}\right]$ and feed it into the least square estimator specified in (19) to update all the active users' channel coefficients $\mathbf{h}^{(n)}$.
8: **end for**
9: Obtain the active user ID set $\mathbf{D}$ based on $\mathbf{C}^{(n_{\max})}$.
10: **return** $\mathbf{D}$, $\hat{\mathbf{h}} = \mathbf{h}^{(n_{\max})}$.

V. Performance Analysis

A. Performance of Layer-by-Layer RM Detection Algorithm

1) Successful Detection Probability: Without loss of generality, we assume that the transmitted RM sequence is $\mathbf{c}^m = \mathbf{1}^{2^m}$ and the elements of the received signal $y_j^m, j = 1, \ldots, 2^m$ are statistically independent. To derive the successful detection

probability of single sequence detection algorithm, we consider the layer-by-layer recovery and the following theorem concludes the correct recovery probability of each layer.

Theorem 2: Under the circumstance that we have recovered $\{\boldsymbol{\alpha}^m, b_m\}, \ldots, \{\boldsymbol{\alpha}^{s+1}, b_{s+1}\}$ correctly, the estimates $\left\{\hat{\boldsymbol{\alpha}}^s, \hat{b}_s\right\}$ are correct with the probability of

$$P^{(s|m,\ldots,s+1)} \geq \max\left\{0, 1-(2^{s-1}-1)\Phi\left(\frac{0-2^{s-1}|h|^2}{\sqrt{2^{s-1}(\tilde{\sigma}_\varepsilon^s)^2}}\right)\right\}, \tag{20}$$

where $(\tilde{\sigma}_\varepsilon^s)^2 = 2|h|^2\frac{N_0}{2^{m-s}} + \left(\frac{N_0}{2^{m-s}}\right)^2$, $2 \leq s \leq m$

and $\Phi(\cdot)$ is the distribution function of the standard normal distribution.

Proof: In the sequel, we prove **Theorem 2** by induction. Firstly, let us consider the case of $s = m$. Recall that $e_j^m, j = 1, \ldots, 2^m$ are i.i.d. random variables following $\mathcal{CN}(0, N_0)$. Given the channel coefficient h, the received signal obeys $y_j^m \sim \mathcal{CN}(h, N_0)$. Then, upon approximating $\tilde{\varepsilon}_j^m$ in (9) by a complex Gaussian variable with mean 0 and variance $(\tilde{\sigma}_\varepsilon^m)^2 \triangleq 2|h|^2 N_0 + N_0^2$, we see that $\tilde{y}_j^m$ follows the distribution $\mathcal{CN}\left(|h|^2, (\tilde{\sigma}_\varepsilon^m)^2\right)$.

Next, the Walsh-Hadamard transformation is performed on $\tilde{\mathbf{y}}^m$ and the result is shown in (10). When $i = 1$, we have

$$V_1^{m-1} = 2^{m-1}\cdot|h|^2 + \sum_{j=1}^{2^{m-1}}\tilde{\varepsilon}_j^m \sim \mathcal{CN}\left(2^{m-1}|h|^2, 2^{m-1}(\tilde{\sigma}_\varepsilon^m)^2\right), \tag{21}$$

while at other locations, i.e., $i = 2, \ldots, 2^{m-1}$, we arrive at:

$$V_i^{m-1} = \sum_{j=1}^{2^{m-1}}(-1)^{(\mathbf{a}_{i-1}^{m-1})^T\overline{\mathbf{a}_{j-1}^{m-1}}}\cdot\tilde{\varepsilon}_j^m \sim \mathcal{CN}\left(0, 2^{m-1}(\tilde{\sigma}_\varepsilon^m)^2\right). \tag{22}$$

Since $\{\boldsymbol{\alpha}^m, b_m\}$ is recovered by searching for the specific element with the largest real component, once there exists some $i \in \{2, \ldots, 2^{m-1}\}$ that results in the event $A_i = \left\{\left(V_1^{m-1}\right)_I - \left(V_i^{m-1}\right)_I < 0\right\}$, the detection will head in the wrong direction. On the basis of (21) and (22), the probability of the event A_i is

$$\Pr(A_i) = \Phi\left(\frac{0-2^{m-1}|h|^2}{\sqrt{2^{m-1}(\tilde{\sigma}_\varepsilon^m)^2}}\right),$$

and the estimates $\left\{\hat{\boldsymbol{\alpha}}^m, \hat{b}_m\right\}$ are correct with the probability

$$P^{(m)} = 1 - \bar{P}^{(m)} = 1 - \Pr\left\{\bigcup_{i=2}^{2^{m-1}} A_i\right\}. \tag{23}$$

Using the classic union bound as the upper bound of the second term in (23) and considering that the probability $P^{(m)}$ cannot be less than 0, we have that

$$\begin{aligned} P^{(m)} &\geq \max\left\{0, 1-\sum_{i=2}^{2^{m-1}}\Pr(A_i)\right\} \\ &= \max\left\{0, 1-(2^{m-1}-1)\Phi\left(\frac{0-2^{m-1}|h|^2}{\sqrt{2^{m-1}(\tilde{\sigma}_\varepsilon^m)^2}}\right)\right\}, \end{aligned} \tag{24}$$

which is in accordance with (20).

Next, assuming that $P^{(s+1|m,\ldots,s+2)}$ also agrees with (20) and the estimated $\left\{\hat{\boldsymbol{\alpha}}^m, \hat{b}_m\right\}, \ldots, \left\{\hat{\boldsymbol{\alpha}}^{s+1}, \hat{b}_{s+1}\right\}$ are correct, y_j^s is calculated as

$$y_j^s = hc_j^s + e_j^s, j = 1, \ldots, 2^s, \tag{25}$$

where $e_j^s = \frac{1}{2}\left(e_j^{s+1} + \hat{v}_j^s \cdot e_{j+2^s}^{s+1}\right)$. With the knowledge that e_j^{s+1} follows the distribution $\mathcal{CN}\left(0, \frac{N_0}{2^{m-s-1}}\right)$, we have that $y_j^s \sim \mathcal{CN}\left(h, \frac{N_0}{2^{m-s}}\right)$. The following derivation is similar to the case of $s = m$ and will not be repeated here due to space limitations. Finally, we can draw the conclusion that the correct detection probability of $\left\{\hat{\boldsymbol{\alpha}}^s, \hat{b}_s\right\}$ is identical to (20). ■

The RM sequence $\mathbf{c}^m$ is detected successfully only if each layer is correctly recovered, hence, its successful detection probability is formulated as

$$P = P^{(m)} \cdot P^{(m-1|m)} \cdots P^{(2|m,m-1,\ldots,3)}.$$

As for the list detection algorithm, it is clear that its performance is expected to be better than that of Algorithm 1, but inferior to that of MLSE.

Remark. Taking a second look at the noise term in (25) , we have:

$$e_j^s = \frac{1}{2}\left(e_j^{s+1} + \hat{v}_j^s \cdot e_{j+2^s}^{s+1}\right) \sim \mathcal{CN}\left(0, \frac{N_0}{2^{m-s}}\right),$$

where we find that the noise level decreases at an exponential rate layer by layer. Thus, if the first few layers are recovered correctly, the subsequent recovery will be correct with a high probability. This finding reveals the necessity to ensure the correctness of the first few layers. Inspired by this, to obtain a better detection performance in the list detection algorithm, we can increase the number of lists for the first few layers, while fixing the total number of paths.

2) Computational Complexity: In the Layer-by-Layer RM detection algorithm, the Walsh-Hadamard transformation (WHT) plays the main role in the computational tasks. Thus, we firstly consider the multiplication operations of the WHT. Thanks to the recursive structure of the Walsh function, a fast WHT can be utilized to reduce the complexity, which takes $\mathcal{O}(s \cdot 2^s)$ multiplications when performed on a length-2^s sequence. Hence, when only the WHT operations are considered, the total complexity of Algorithm 1 is $\mathcal{O}((m-2)2^m)$. By contrast, the algorithm given in [16] exhibits a complexity order up to $\mathcal{O}\left(m^2 \cdot 2^m\right)$. This is due to the fact that the algorithm in [16] needs a length-2^m fast WHT to recover

each column of the matrix $\mathbf{P}^m$, while in Algorithm 1, the fast WHT is performed on $\tilde{\mathbf{y}}^s, s = m, m-1, \ldots, 2$, whose length is decreasing exponentially layer by layer. Hence, our proposed Algorithm 1 has significant advantage on complexity. If we further take the multiplications in Eq. (9) and Eq. (12) into account, the complexity of Algorithm 1 and the algorithm in [16] is $\mathcal{O}\left[(m+1)\cdot 2^m\right]$ and $\mathcal{O}\left[(2m^2+m+2)\cdot 2^m\right]$, respectively.

It is intuitive that the list detection algorithm improves the performance at the cost of an increased complexity. For the $(\mathbf{L}, F)$ list detection algorithm, the number of multiplications is given by

$$\begin{aligned} N_M(\mathbf{L},F) =& (m+2)2^{m-1} \\ &+ \sum_{s_1=m-F+1}^{m} \left(\prod_{s_2=s_1}^{m} L_{s_2}\right)(s_1+1)2^{s_1-2} \\ &+ \left(\prod_{s_1=m-F+1}^{m} L_{s_1}\right)\sum_{s_2=4}^{m-F}(s_2+1)2^{s_2-2}. \end{aligned}$$

Zhang et al. [17] proposed a shuffled version of the algorithm in [16] , which aims to improve the performance by changing the recovery order of the columns of the matrix $\mathbf{P}^m$. When the number of shuffles is S, the complexity of this enhanced algorithm equals to $\mathcal{O}\left[(2m^2+m+2)\cdot 2^m S\right]$.

Moreover, it is worth to mention the channel estimation complexity difference between these algorithms. When executed all the way to the last step, Algorithm 1 can directly estimate the channel coefficients according to (15). While the algorithms proposed in [16] and [17] cannot compute the channel coefficients and an extra least square (LS) estimation problem has to be solved, which incurs additional computational complexity.

B. Performance of the Iterative RM Detection and Channel Estimation Algorithm

For ease of clarification, we analyze the performance in the case of $K = 2$. At Stage $(1,1)$, the input of Algorithm 1 is exactly the received signal, i.e.,

$$y_j^{(1,1)} = h_1 \cdot c_{1,j}^m + h_2 \cdot c_{2,j}^m + e_j^m, j = 1, \ldots, 2^m. \tag{26}$$

After splitting it into a pair of partial sequences and computing their element-wise conjugate multiplication, we have

$$\begin{aligned} \tilde{y}_j^m =& |h_1|^2 \cdot v_{1,j}^{m-1} + |h_2|^2 \cdot v_{2,j}^{m-1} + h_1^* h_2 c_{1,j}^{m-1} c_{2,j}^{m-1} v_{1,j}^{m-1} \\ &+ h_1 h_2^* c_{1,j}^{m-1} c_{2,j}^{m-1} v_{2,j}^{m-1} + \tilde{\varepsilon}_j^m, j = 1, \ldots, 2^{m-1}, \end{aligned} \tag{27}$$

where $\tilde{\varepsilon}_j^m$ follows the distribution $\mathcal{CN}\left(0, (\tilde{\sigma}_\varepsilon^m)^2\right)$ and $(\tilde{\sigma}_\varepsilon^m)^2 \triangleq 2\left(|h_1|^2 + |h_2|^2\right)N_0 + N_0^2$. Then, the Walsh-Hadamard transformation is applied to (27) for obtaining the vector $\mathbf{V}^{m-1}$. Under the assumption that $|h_1|^2 \geq |h_2|^2$, $\boldsymbol{\alpha}_1^m$ is recovered correctly if and only if $\left(V_{\langle \boldsymbol{\alpha}_1^m \rangle}^{m-1}\right)_I > \left(V_i^{m-1}\right)_I$ is satisfied for every $i \in \{1, \ldots, 2^{m-1}\} \backslash \langle \boldsymbol{\alpha}_1^m \rangle$, where $\langle \boldsymbol{\alpha}_1^m \rangle$ is the location index that satisfies $\boldsymbol{\alpha}_1^m = \mathbf{a}_{\langle \boldsymbol{\alpha}_1^m \rangle - 1}^{m-1}$.

Firstly, in the special case of $\boldsymbol{\alpha}_1^m = \boldsymbol{\alpha}_2^m$, the real component of the element at the location $\langle \boldsymbol{\alpha}_1^m \rangle$ turns out to be

$$\begin{aligned} &\left(V_{\langle \boldsymbol{\alpha}_1^m \rangle}^{m-1}\right)_I \\ &= 2^{m-1}\left[(-1)^{b_{1,m}} \cdot |h_1|^2 + (-1)^{b_{2,m}} \cdot |h_2|^2\right] \\ &+ \left[(-1)^{b_{1,m}} + (-1)^{b_{2,m}}\right](h_1^* h_2)_I \sum_{j=1}^{2^{m-1}} \left(c_{1,j}^{m-1} c_{2,j}^{m-1}\right) \\ &+ \left((\varepsilon_{1,j}^m)'\right)_I. \end{aligned} \tag{28}$$

The benefit of our proposed mapping rule manifests itself here. It is obvious from (28) that if $b_{1,m} \neq b_{2,m}$, the component $\left(V_{\langle \boldsymbol{\alpha}_1^m \rangle}^{m-1}\right)_I$ will decrease to $2^{m-1}\left(|h_1|^2 - |h_2|^2\right) + \left((\varepsilon_{1,j}^m)'\right)_I$ and there is a substantial chance that it will become smaller than the other interference components, which results in the incorrect recovery of $\boldsymbol{\alpha}_1^m$. While our constraint in (2) ensures that $b_{1,m} = b_{2,m}$ in this case and the magnitude of $\left(V_{\langle \boldsymbol{\alpha}_1^m \rangle}^{m-1}\right)_I$ is reinforced to be

$$\begin{aligned} \left(V_{\langle \boldsymbol{\alpha}_1^m \rangle}^{m-1}\right)_I =& 2^{m-1}\left(|h_1|^2 + |h_2|^2\right) \\ &+ 2(h_1^* h_2)_I \sum_{j=1}^{2^{m-1}} \left(c_{1,j}^{m-1} c_{2,j}^{m-1}\right) + \left((\varepsilon_{1,j}^m)'\right)_I, \end{aligned} \tag{29}$$

which is expected to improve the detection performance.

Next, in the case of $\boldsymbol{\alpha}_1^m \neq \boldsymbol{\alpha}_2^m$, the real component of the element $V_{\langle \boldsymbol{\alpha}_1^m \rangle}^{m-1}$ is specified as

$$\begin{aligned} &\left(V_{\langle \boldsymbol{\alpha}_1^m \rangle}^{m-1}\right)_I = 2^{m-1}|h_1|^2 \\ &+ (h_1^* h_2)_I \sum_{j=1}^{2^{m-1}} \left(c_{1,j}^{m-1} c_{2,j}^{m-1} + c_{1,j}^{m-1} v_{1,j}^{m-1} \cdot c_{2,j}^{m-1} v_{2,j}^{m-1}\right) \\ &+ \left((\varepsilon_{1,j}^m)'\right)_I, \end{aligned} \tag{30}$$

where $(\varepsilon_{1,j}^m)' \triangleq \sum_{j=1}^{2^{m-1}} \tilde{\varepsilon}_j^m \cdot v_{1,j}^{m-1} \sim \mathcal{CN}\left(0, 2^{m-1}(\tilde{\sigma}_\varepsilon^m)^2\right)$. The second term of (30) happens to be the inner product of $\mathbf{c}_1^m$ and $\mathbf{c}_2^m$, which is denoted as $\chi_{1,2}^m$. Then, Eq. (30) can be recast as

$$\left(V_{\langle \boldsymbol{\alpha}_1^m \rangle}^{m-1}\right)_I = 2^{m-1}|h_1|^2 + (h_1^* h_2)_I \cdot \chi_{1,2}^m + \left((\varepsilon_{1,j}^m)'\right)_I. \tag{31}$$

Similarly, we have that

$$\left(V_{\langle \boldsymbol{\alpha}_2^m \rangle}^{m-1}\right)_I = 2^{m-1}|h_2|^2 + (h_1^* h_2)_I \cdot \chi_{1,2}^m + \left((\varepsilon_{2,j}^m)'\right)_I.$$

While at the location i $(i \neq \langle \boldsymbol{\alpha}_1^m \rangle$ and $i \neq \langle \boldsymbol{\alpha}_2^m \rangle)$, the real component is expressed as

$$\begin{aligned} &\left(V_i^{m-1}\right)_I = (h_1^* h_2)_I \\ &\cdot \sum_{j=1}^{2^{m-1}} \left(c_{1,j}^{m-1} c_{2,j}^{m-1} \cdot v_{1,j}^{m-1} \cdot t_{i,j}^{m-1} + c_{1,j}^{m-1} c_{2,j}^{m-1} \cdot v_{2,j}^{m-1} \cdot t_{i,j}^{m-1}\right) \\ &+ \left((\varepsilon_{i,j}^m)'\right)_I, \end{aligned}$$

recalling that $t_{i,j}^{m-1}$ is the element of the Hadamard matrix $\mathbf{T}^{m-1}$. To proceed, we recast the first term as follows:

$$\sum_{j=1}^{2^{m-1}} \left(c_{1,j}^{m-1} c_{2,j}^{m-1} v_{1,j}^{m-1} t_{i,j}^{m-1} + c_{1,j}^{m-1} c_{2,j}^{m-1} v_{2,j}^{m-1} t_{i,j}^{m-1}\right)$$

$$
\begin{aligned}
&= \sum_{j=1}^{2^{m-1}} \left(c_{1,j}^{m-1} c_{2,j}^{m-1} + c_{1,j}^{m-1} v_{1,j}^{m-1} c_{2,j}^{m-1} t_{i,j}^{m-1}\right) \\
&\quad - \sum_{j=1}^{2^{m-1}} \left(c_{1,j}^{m-1} c_{2,j}^{m-1}\right) \\
&\quad + \sum_{j=1}^{2^{m-1}} \left(c_{1,j}^{m-1} c_{2,j}^{m-1} + c_{1,j}^{m-1} t_{i,j}^{m-1} c_{2,j}^{m-1} v_{2,j}^{m-1}\right) \\
&\quad - \sum_{j=1}^{2^{m-1}} \left(c_{1,j}^{m-1} c_{2,j}^{m-1}\right) \\
&= \chi_{1,(2,i)}^{m} - \chi_{1,2}^{m-1} + \chi_{(1,i),2}^{m} - \chi_{1,2}^{m-1},
\end{aligned}
$$

where $\chi_{1,(2,i)}^{m}$ denotes the inner product of $\mathbf{c}_1^m$ and the RM sequence generated by the matrix-vector expressed as

$$
\begin{aligned}
\mathbf{P}_{(2,i)}^{m} &= \begin{bmatrix} 0 & \left(\mathbf{a}_{i-1}^{m-1}\right)^T \\ \mathbf{a}_{i-1}^{m-1} & \mathbf{P}_2^{m-1} \end{bmatrix}, \\
\mathbf{b}_{(2,i)}^{m} &= \left[a_{i-1,1}^{m-1} \oplus \cdots \oplus a_{i-1,m-1}^{m-1}, \mathbf{b}_2^{m-1}\right],
\end{aligned}
$$

and $\chi_{(1,i),2}^{m}$ is obtained in the similar way.

It is obvious that the inner products of the RM sequences play an important role in determining the performance of multi-sequence detection. The authors of [16] reveal that the distribution of the inner products of second-order RM sequences is symmetric. Moreover, if two distinct matrices $\mathbf{P}_1^m$ and $\mathbf{P}_2^m$ satisfy that

$$
\text{rank}(\mathbf{P}_1^m - \mathbf{P}_2^m) = 2r, r = 1, \ldots, [m/2],
$$

then the magnitude of the inner product of their respective generated RM sequences is either 2^{m-r} or 0. This inspires us to restrict the matrices, for example, to be taken from the $DG(m, 2r)$ set [18], for further improving the attainable performance of RM detection.

Then, on the premise that $\{\boldsymbol{\alpha}_1^m, b_{m,1}\}$ is correctly recovered, the sequence $\mathbf{y}^{m-1}$ is given as

$$
y_j^{m-1} = h_1 \cdot c_{1,j}^{m-1} + \frac{1}{2} h_2 c_{2,j}^{m-1} \left(1 + v_{1,j}^{m-1} \cdot v_{2,j}^{m-1}\right) + e_j^{m-1},
$$

where $e_j^{m-1} = \frac{1}{2}\left(e_j^m + v_{1,j}^{m-1} \cdot e_{j+2^{m-1}}^m\right)$. Combined with (26), it is found that for the first active user, the signal-to-interference-and-noise-ratio (SINR) is enhanced. Furthermore, as the algorithm proceeds, the SINR of $\mathbf{y}^s$, which is denoted as γ^s, doubles layer by layer, specifically,

$$
\gamma^s = \frac{2^{m-s}|h_1|^2}{|h_2|^2 + N_0}, s = m, \ldots, 1. \tag{32}
$$

This finding further underlines the necessity of ensuring the correctness of the first few layers.

At the end of Stage $(1, 1)$, the estimates $\mathbf{c}_1^{(1)}$ and $h_1^{(1)}$ are obtained. Then we enter Stage $(1, 2)$. At the beginning of Stage $(1, 2)$, the residual signal is calculated by

$$
y_j^{(1,2)} = h_2 c_{2,j}^m + \left(h_1 c_{1,j}^m - h_1^{(1)} c_{1,j}^{(1)}\right) + e_j^m, j = 1, \ldots, 2^m, \tag{33}
$$

which is then input to Algorithm 1. The accuracy of both $\mathbf{c}_1^{(1)}$ and $h_1^{(1)}$ affects the performance at this stage. If the sequence $\mathbf{c}_1^{(1)}$ is correct, the impact of the channel estimation error (CEE) is relatively small. On one hand, in the case that $\boldsymbol{\alpha}_1^s, s = m, \ldots, 2$ are all correctly recovered, the SINR of the last layer is up to $\frac{2^{m-1}|h_1|^2}{|h_2|^2+N_0}$ according to (32), which ensures that the CEE will not be too high. On the other hand, the LS channel estimator at the end of the first iteration is capable of further improving the accuracy of $h_1^{(1)}$, and the second user's RM sequence can be updated in the second iteration.

Whereas if the sequence $\mathbf{c}_1^{(1)}$ is incorrect, (33) is equivalent to adding an extra sequence to the original received signal, which further lower the SINR. In [16] and [17], a traditional SIC-based method is used for multi-sequence detection . Explicitly, before detecting the next sequence, all the sequences obtained so far are subtracted from the received signal. In this scheme, the impact of any incorrectly detected sequences always persists. By contrast, the iterative method adopted in our proposed algorithm deals with this problem in a better way. For example, if the detected sequence $\mathbf{c}_1^{(1)}$ is wrong but $\mathbf{c}_2^{(1)}$ happens to be right, then at Stage $(2, 1)$, after subtracting the MAI term, the signal input to Algorithm 1 is equal to

$$
h_1 c_{1,j}^m + \left(h_2 c_{2,j}^m - h_2^{(1)} c_{2,j}^{(1)}\right) + e_j^m, j = 1, \ldots, 2^m,
$$

which is free from the impact of the incorrect result $\mathbf{c}_1^{(1)}$ and $h_1^{(1)}$. Moreover, the algorithms in [16] and [17] need to perform LS channel estimation after each sequence detection, while our proposed algorithm executes LS channel estimation only once in an iteration, which further reduces the complexity.

Finally, the properties of our proposed algorithms are summarized as follows:

- As a benefit of the proposed mapping method, the channel coefficients can be estimated directly according to (15), which leads to a reduced computational complexity without any performance loss.
- In the case where multiple active users possess the same $\boldsymbol{\alpha}_k^s$, our proposed mapping rule ensures that the corresponding result of the WHT is reinforced like in (V-B). As such, the successful recovery probability of $\boldsymbol{\alpha}_k^s$ is expected to be enhanced.
- By exploiting the nested structure of RM sequences, the matrix-vector pair is recovered layer by layer. During this process, the interference and noise level is found to keep on decreasing on the premise that $\{\boldsymbol{\alpha}^s, b_s\}$ are recovered correctly. On one hand, this ensures the accuracy of channel estimation based on (15). On the other hand, this reminds us of the importance of improving the recovery reliability of the first few layers. Inspired by this, we can adjust the parameters of the list detection algorithm for enhancing the performance.
- The inner products of RM sequences have a significant influence on the performance of multi-sequence detection. Motivated by this, we can further restrict the choice of the matrices $\mathbf{P}^m$ to enhance detection capability. Intuitively, this is achieved at the cost of shrinking the user space size.
- In contrast to the common SIC-based detection, the iterative detection algorithm utilized in this paper is capable

of reducing the impact of incorrectly detected sequences, thus it has a better capability to deal with the error propagation problem of SIC.

VI. Simulation Results

In this section, we evaluate the performance of the proposed algorithms in terms of their sequence detection accuracy, channel estimation accuracy and computational cost. For ease of exposition, Algorithm 1 and its list detection version are denoted as "*RM_LLD*" and "*list RM_LLD*", respectively. Moreover, the algorithm in [16] is used as the benchmark. Since it is based on the application of shift-and-multiply to the received signal, we referred to it and its shuffled version proposed in [17] as "*RM_SMD*" and "*RM_SMD with shuffle*", respectively. In the sequel, the successful detection probability is defined as the average ratio between the number of successfully detected active users and the total number of active users in the system. The channel estimation error is quantified by the mean square error of channel estimates of successfully detected active users.

A. Performance of the Layer-by-Layer RM Detection Algorithm

Firstly, we examine the performance of the proposed algorithms in the single-sequence scenario. The simulation parameters are listed in Table. I.

TABLE I
Simulation Parameters of Single-Sequence Detection.

Sequence Length	2^8
SNR	$-10 \sim 5$dB
Shuffle Times in "*RM_SMD with shuffle*"	4
$(\mathbf{L}, F)$ of "*list RM_LLD*"	$([2,2],2)$
Simulation Times	2000

Fig. 7 illustrates the successful detection probability versus the signal to noise ratio (SNR). It can be observed that *RM_SMD* faces severe performance degradation in the low-SNR scenario, while *RM_LLD* exhibits robustness against the SNR, thus confirming the superior detection performance of the proposed algorithm. As expected, both the shuffling operation and the list detection approach are capable of improving the detection performance. Even though both *RM_SMD with 4 shuffles* and *list RM_LLD with* $([2,2],2)$ choose the optimal result from 4 candidates, it is shown in Fig. 7 that *list RM_LLD with* $([2,2],2)$ is superior. Moreover, the theoretical results obtained in Section V.A is also depicted in Fig. 7. Since the union bound is utilized in the analysis, this result can act as the lower bound. The simulation result of *RM_LLD* and the theoretical result consistently show that when the SNR is higher than -4dB, the RM sequence can be successfully detected with a probability close to 1.

The computational complexity of these four algorithms associated with $m = 8$ is compared in Table II. Here we quantify the complexity by the number of multiplication operations. Observe that the proposed *RM_LLD* impose a much

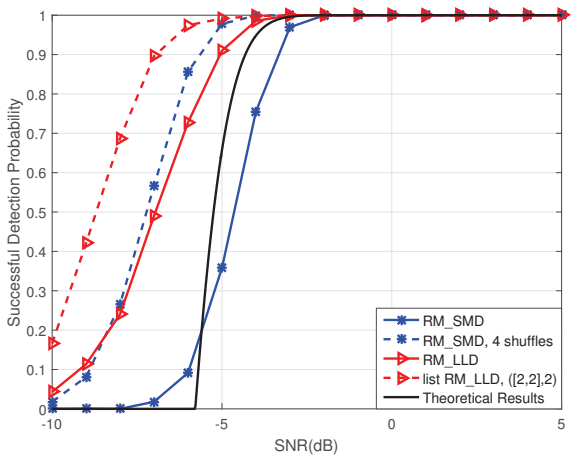


Fig. 7. The successful reconstruction probability versus the signal to noise ratio (SNR) when $m = 8$.

lower complexity than *RM_SMD*. It is intuitive that both the list detection approach and the shuffle operation improve the performance at the cost of an increased complexity. However, the complexity of *RM_SMD with shuffles* increases linearly with the number of shuffles, while the complexity growth rate of *list RM_LLD* is much more gradual.

TABLE II
Comparison of different algorithms in terms of computational complexity in the case of $m = 8$.

Algorithm	Computational Complexity
RM_LLD	2304
list RM_LLD, $([2,2],2)$	3820
RM_SMD	35328
RM_SMD, 4 shuffles	141312

Finally, we verify the feasibility of estimating the channel coefficients according to (15). Fig. 8 depicts the channel estimation errors calculated both from Eq. (15) and by the least square (LS) estimator. It can be seen from Fig. 8 that the channel estimates calculated by (15) exhibit the same accuracy as the results obtained by the LS estimator. However, it is obvious that the complexity of (15) is much lower, which further validates the superiority of our proposed algorithm.

B. Performance of the iterative RM Detection and Channel Estimation Algorithm

Next, we evaluate the performance of our algorithms in the multi-sequence scenario and the corresponding parameters are summarized in Table. III.

The successful detection probability versus the number of active users in the case of $m = 8$ and $m = 10$ is depicted in Fig. 9. It can be observed that the *RM_SMD* curve drops rapidly when increasing the number of active users, while *RM_LLD* shows a significant advantage over it. As in the single sequence scenario, both "*RM_SMD with shuffle*" and "*list RM_LLD*" improve the successful detection

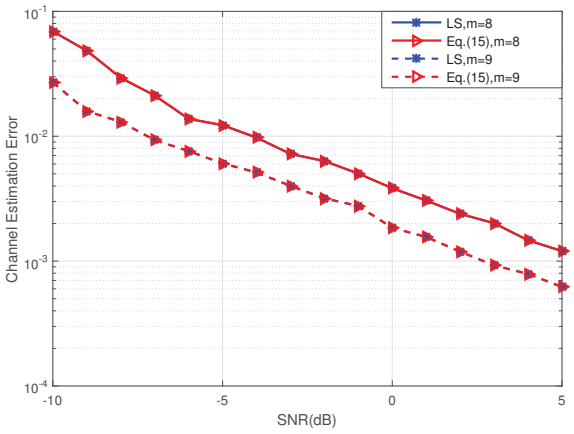
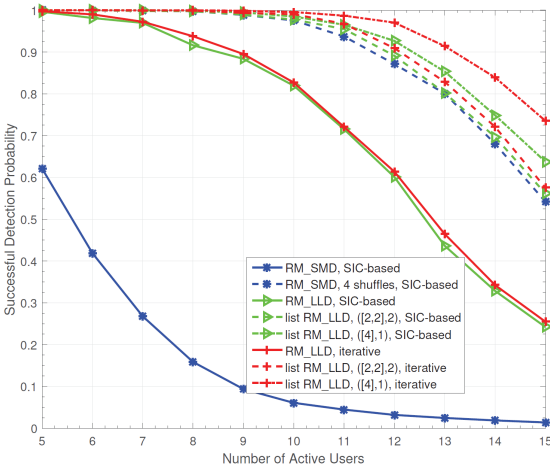
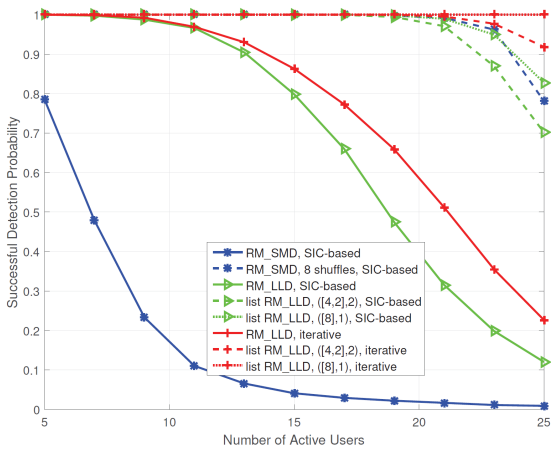


Fig. 8. The channel coefficient estimation error versus SNR.



(a) $m = 8$



(b) $m = 10$

Fig. 9. The successful detection probability versus the number of active users in the system.

TABLE III
SIMULATION PARAMETERS OF MULTI-SEQUENCE DETECTION.

Sequence Length	2^8, 2^{10}
SNR	20dB
# of iterations in "*RM_LLD, iterative*"	5
False Alarm Rate	$\leq 10^{-7}$
Shuffle Times in "*RM_SMD with shuffle*"	4 for length-2^8 sequences; 8 for length-2^{10} sequences
$(\mathbf{L}, F)$ of "*list RM_LLD*"	$([2,2],2)$ and $([4],1)$ for length-2^8 sequences; $([4,2],2)$ and $([8],1)$ for length-2^{10} sequences
Simulation Times	2000

probability. Comparing the successful detection probability of "*list RM_LLD,* $([2,2],2)$" and "*list RM_LLD,* $([4],1)$", we find that with the total number of paths fixed, the performance can be beneficially improved by increasing the number of lists in the first layer. This result further validates the importance of ensuring the reliable recovery of the first few layers in the proposed RM detection algorithm. Moreover, it can be seen from Fig. 9 that the iterative method adopted by our proposed algorithm is superior to the traditional SIC-based multi-sequence detection. This performance improvement is an explicit benefit of the mutual information exchange between the RM detector and the channel estimator as well as of the capability to eliminate the impact of incorrectly detected sequences. Combining Fig. 9(a) and (b), it is also clear that the detection capability is enhanced by increasing the length of sequences. Fig. 10 compares the channel estimate

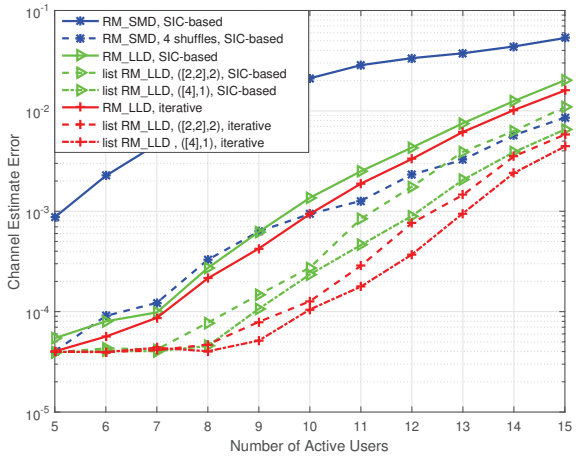


Fig. 10. The channel coefficient estimation error versus the number of active users in the system in the case of $m = 8$.

error of different algorithms in the multi-sequence scenario when $m = 8$. An insightful observation is that although the successful detection probability of "*list RM_LLD, SIC-based*" and "*list RM_LLD, iterative*" seems similar in Fig.

9 (a), the channel estimates of "*list RM_LLD, iterative*" are obviously more accurate than those of "*list RM_LLD, SIC-based*". Furthermore, the advantage of the iterative method becomes more prominent for "*list RM_LLD,* $([2,2],2)$" and "*list RM_LLD,* $([4],1)$". Based on these results, we may conclude that the proposed iterative RM detection and channel estimation algorithm beneficially improves the user detection and channel estimation performance.

Fig. 11 depicts the convergence of the iterative RM detection and channel estimation algorithm for $m = 8$. It can be seen that the convergence rate is related to the number of active users in the system. In the case of $K \leq 10$, the algorithm can converge within 5 iterations. Although it needs more iterations to converge when the number of active users increases, the successful detection probability keeps rising before achieving the convergence. Hence, the value of $n_{\max}$ can be set according to the practical tradeoff between the detection capability and the complexity.

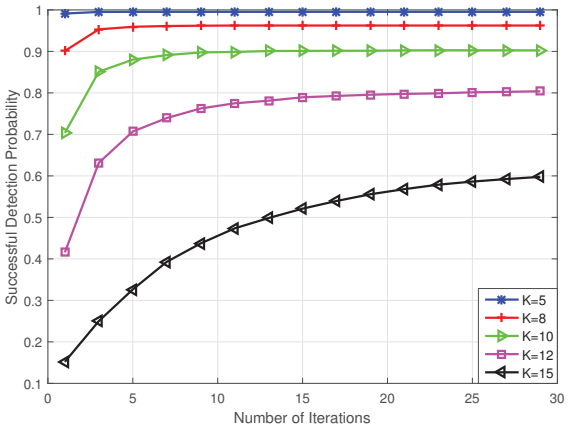


Fig. 11. The convergence of the iterative RM detection and channel estimation algorithm.

VII. Conclusions

In this paper, we utilize RM sequences for user identification and channel estimation to meet the massive connectivity and low latency requirements of mMTC scenarios. A novel mapping rule was conceived for improving RM detection capability and for simplifying the channel estimation with no loss of accuracy. At the receiver, a layer-by-layer RM detection algorithm and an enhanced algorithm were invoked for single sequence detection. They are based on the nested structure of RM sequences, which is discovered to reveal the deeper relationship between their sub-sequences. Then an iterative RM detection and channel estimation algorithm was also designed for the case where multiple active users co-exist in the system. From the simulation results, our proposed algorithms have a significant advantage over existing ones in terms of detection performance and computational complexity.

References

[1] ITU-R Rec. M.2083-0, IMT Vision - Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond, Sept. 2015.

[2] Huawei, 5G Network Architecture-A High-Level Perspective, 2016, white paper at Huawei.com.

[3] 3GPP technical specifications 36.321. Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC). Available at: www.3gpp.org.

[4] 3GPP technical specifications 36.211. Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation. Available at: www.3gpp.org.

[5] G. Durisi, T. Koch, and P. Popovski, "Towards Massive, Ultra-Reliable, and Low-Latency Wireless Communication with Short Packets," *Proc. IEEE*, vol. 104, no. 9, pp. 1711-1726, 2016.

[6] H. S. Jang, S. M. Kim, K. S. Ko, J. Cha, and D. K. Sung, "Spatial group based random access for M2M communications," *IEEE Commun. Lett.*, vol. 18, no. 6, pp. 961-964, 2014.

[7] S. Kim, H. Kim, H. Noh, Y. Kim, and D. Hongy, "Novel Transceiver Architecture for an Asynchronous Grant-free NOMA System," unpublished.

[8] N. K. Pratas, H. Thomsen, C. Stefanovic, and P. Popovski, "Code-Expanded Random Access for Machine-Type Communications," in *IEEE Globecom Workshops (GC Wkshps)*, Anaheim, 2012, pp. 1681-1686.

[9] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203-4215, 2005.

[10] L. Liu and W. Yu, "Massive Connectivity With Massive MIMO-Part I: Device Activity Detection and Channel Estimation," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2933-2946, 2018.

[11] G. Wunder, P. Jung, and M. Ramadan, "Compressive Random Access Using a Common Overloaded Control Channel," in *IEEE Globecom Workshops (GC Wkshps)*, San Diego, 2015, pp. 1-6.

[12] H. F. Schepker, C. Bockelmann, and A. Dekorsy, "Exploiting Sparsity in Channel and Data Estimation for Sporadic Multi-User Communication," in *Tenth Int. Symp. Wireless Commun. Syst.*, Ilmenau, 2013, pp. 1-5.

[13] G. Hannak, M. Mayer, A. Jung, G. Matz, and N. Goertz, "Joint channel estimation and activity detection for multiuser communication systems," in *IEEE Int. Conf. Commun. Workshop (ICCW)*, London, 2015, pp. 2086-2091.

[14] Z. Chen and W. Yu, "Massive device activity detection by approximate message passing," in *IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, New Orleans, 2017, pp. 3514-3518.

[15] N. K. Pratas, C. Stefanovic, G. C. Madueno and P. Popovski, "Random Access for Machine-Type Communication Based on Bloom Filtering," in *IEEE Global Commun. Conf. (GLOBECOM)*, Washington, 2016, pp. 1-7.

[16] S. D. Howard, A. R. Calderbank, and S. J. Searle, "A fast reconstruction algorithm for deterministic compressive sensing using second order reed-muller codes," in *42nd Annu. Conf. Inf. Sci. and Sys.*, Princeton, 2008, pp. 11-15.

[17] H. Zhang, R. Li, J. Wang, Y. Chen and Z. Zhang, "Reed-Muller Sequences for 5G Grant-Free Massive Access," in *IEEE Global Commun. Conf.*, Singapore, 2017, pp. 1-7.

[18] A. Calderbank and S. Jafarpour, "Reed Muller sensing matrices and the LASSO," in *Int. Conf. Sequences Their Applicat. (SETA)*, Paris, 2010, pp. 442-463.

Jue Wang received the B.S.Eng. degree in communication engineering from the Department of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China, in 2016. She is currently working towards her Ph.D. degree in information and communication engineering at Zhejiang University, Hangzhou, China. Her research interests mainly include signal processing, massive access, massive MIMO and machine learning.

Zhaoyang Zhang (M'02) received his Ph.D. degree from Zhejiang University, Hangzhou, China, in 1998, where he is currently a Qiushi Distinguished Professor. His current research interests are mainly focused on the fundamental aspects of wireless communications and networking, such as information theory and coding, network signal processing and distributed learning, AI-empowered communications and networking, synergistic sensing, communication and compuation, etc. He has published more than 300 peer-reviewed international journal and conference papers, including 6 conference best papers. He was awarded the National Natural Science Fund for Outstanding Young Scholars by NSFC in 2017. He is serving as Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS and IET COMMUNICATIONS, etc, and has served as General Chair, TPC Co-Chair or Symposium Co-Chair for WCSP 2013/2018, Globecom 2014 Wireless Communications Symposium, and VTC-Spring 2017 Workshop HMWC, etc.

Lajos Hanzo (http://www-mobile.ecs.soton.ac.uk) FREng, F'04, FIET, Fellow of EURASIP, received his 5-year degree in electronics in 1976 and his doctorate in 1983 from the Technical University of Budapest. In 2009 he was awarded an honorary doctorate by the Technical University of Budapest and in 2015 by the University of Edinburgh. In 2016 he was admitted to the Hungarian Academy of Science. During his 40-year career in telecommunications he has held various research and academic posts in Hungary, Germany and the UK. Since 1986 he has been with the School of Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He has successfully supervised 119 PhD students, co-authored 18 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, published 1800+ research contributions at IEEE Xplore, acted both as TPC and General Chair of IEEE conferences, presented keynote lectures and has been awarded a number of distinctions. Currently he is directing a 60-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) UK, the European Research Council's Advanced Fellow Grant and the Royal Society's Wolfson Research Merit Award. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a Governor of the IEEE ComSoc and VTS. He is a former Editor-in-Chief of the IEEE Press and a former Chaired Professor also at Tsinghua University, Beijing.