Composite Preambles Based on Differential Phase Rotations for Grant-free Random Access Systems

Yang Wang, Student Member, IEEE, Wenjun Xu, Senior Member, IEEE, Markku Juntti, Fellow, IEEE, Jiaru Lin, Member, IEEE, Miao Pan, Senior Member, IEEE

Abstract—With the advantages of low signaling overhead and latency, grant-free random access (GFRA) becomes a promising technology for supporting massive machine-type communications (mMTC), but poses new challenges for active user detection (AUD) and channel estimation (CE), whose performance mainly depends on the preamble detection. In this paper, we design the composite preamble based on differential phase rotations by aggregating orthogonal Zadoff-Chu (ZC) sequences and multiple root ZC sequences with differential phase rotations to reduce the probability of preamble collisions, thereby improving the performance of AUD and CE. In particular, differential phase rotations extend the preamble set size so that users colliding in orthogonal sequences can be distinguished by phase rotations. In addition, it also reduces non-orthogonal interference and thus reduces CE errors. The preamble detection algorithm and CE scheme are proposed, along with the theoretical analysis of AUD and CE performance to verify the effectiveness of the designed preamble. In addition, the proposed preamble is extended to combine phase rotations with cyclic shifts to further enlarge the preamble set size with low non-orthogonality. Simulation results show that the proposed composite preamble outperforms existing preambles in terms of the probability of detection and CE accuracy.

Index Terms—Composite preambles, massive machine-type communications, preamble collision, user detection.

I. INTRODUCTION

The past few decades have witnessed the proliferation of Internet-of-Things (IoT) systems, especially the recent rapid rise of machine-type communications (MTC) devices that access the Internet through wireless networks, such as vehicles, sensors, household appliances and smart medical devices [1]. Statistics show that MTC devices are expected to account for more than 50% of devices connected to communication networks by 2023 [2], [3]. In line with this trend, ITU has identified massive machine-type communications (mMTC) as one of the three representative scenarios for 5G wireless systems [4].

Typically, mMTC is characterized by massive connections, sporadic transmission, and short data packets. In such cases, the existing grant-based random access (RA) protocol is not desirable due to the excessive latency and signaling overhead. Therefore, the grant-free RA (GFRA) protocol is proposed to accommodate mMTC, which allows devices to transmit packets consisting of preamble and data at any time without getting a grant from the base station (BS), and thus reducing the signaling overhead and latency [5]. Moreover, to support massive connectivity, grant-free access is generally used in non-orthogonal multiple access (NOMA) and multiple-input and multiple-output (MIMO) networks [5]–[7].

In GFRA systems, since the BS has no information about user scheduling, the active user detection (AUD) and channel estimation (CE) should be performed first by processing preambles to facilitate the subsequent data detection. The existing schemes mostly assume that the user accesses the channel with a unique preamble, and then detect the presence of preambles to distinguish between active and inactive devices [8]–[11]. However, it is difficult to guarantee the uniqueness of preambles. On the one hand, as the choice of preambles is random and independent, it is highly likely that two users choose the same preamble, i.e., a preamble collision occurs. On the other hand, the number of available preambles (i.e., preamble set size) is limited due to the limited preamble length, and the probability of preamble collisions will be increased with the growth of the number of MTC devices. In the case of collisions, the accuracy of the AUD is degraded, resulting in a significant increase in the CE and data detection errors. Therefore, preamble collisions are the bottleneck of performance.

Several orthogonal sequence-based schemes have been proposed to resolve the problem. It has been demonstrated that when the number of antennas is sufficiently large, orthogonal preambles enable accurate AUD and CE, and the performance of GFRA systems depends mainly on the preamble length [12]. To detect users suffering from preamble collisions, a non-orthogonal random access (NORA) scheme has been proposed, which distinguishes collided users based on the difference in times of arrival (ToA) [13]. However, the collided users cannot be detected when their ToA are close to each other. Then, more attention is paid to assembling orthogonal preambles in
non-orthogonal manners to enlarge the preamble pool and thus avoiding preamble collisions.

Protocols that send preambles with different patterns have been proposed, including sending preambles by hopping over multiple slots [14] and inserting preambles in designated slots [15]. These patterns mitigate collisions, but also pose challenges for AUD and CE. Inspired by [14] and [15], a super preamble (S-preamble) that sends multiple orthogonal preambles in succession has been proposed in [16], which enlarges the preamble set size at the cost of excessively long preamble transmission phases. Subsequently, an improved S-preamble has been proposed in [17] to expand the preamble pool without increasing the preamble length, which divides a transmission phase into several consecutive subphases and sends a shorter orthogonal preamble in each sub-phase. Nevertheless, it cannot be ignored that the performance of CE degrades due to the non-orthogonality of S-preambles, and it is shown that preambles with subphases not always outperform those without them [18]. Therefore, preambles without subphases have been proposed. Specifically, non-orthogonal preambles formed by the union of two orthogonal preamble sets have been proposed in [19] and [20], where the selection of preamble sets obeys predefined rules. Preambles consisting of fixed preambles and random sequences have been proposed in [21], where the random sequences have been used for data scrambling. Note that the preambles in [21] are carefully designed for two pre-defined categories of users, and there are one-to-one mapping relationships between fixed preambles and random sequences. The predefined rules and one-to-one mappings reduce the non-orthogonal interference, but also limit the preamble set size to some extent. Besides, preambles formed by the sum of randomly selected orthogonal sequences have also been proposed, where sequences are transmitted at different power levels [22], [23]. The power levels expand the preamble pool but reduce the accuracy of the AUD.

In addition to assembling orthogonal sequences, a more straightforward approach is to use non-orthogonal sequences as preambles, which naturally increases the preamble set size despite decreasing the accuracy of the detection. In detail, preambles based on Bernoulli sequences [24], m-sequences [25], Gaussian sequences [7] and multiple root Zadoff-Chu (ZC) sequences [6] have been designed, among which Gaussian sequences and multiple root ZC sequences are most widely adopted. The performance of the two sequences is compared in [26], and it is shown that multiple root ZC sequences outperform Gaussian sequences in most cases.

Based on the existing schemes, it can be concluded that preambles based on either orthogonal or non-orthogonal sequences are not competent since orthogonal sequences are not sufficient to support massive connections and non-orthogonal sequences have been criticized for having large interference. Therefore, the key problem becomes how to expand the size of the preamble pool at the cost of lower non-orthogonal interference. Different from existing schemes, we propose the composite preamble by aggregating orthogonal and differential phase rotation-based nonorthogonal sequences to solve the problem. Specifically, the proposed preamble is designed as an orthogonal sequence followed by a non-orthogonal sequence, where the orthogonal sequences are used to guarantee the performance of AUD and CE of users without orthogonal preamble collisions as well as to reduce the CE errors of collided users. As for the non-orthogonal sequences, they are generated by multiple root ZC sequences with differential phase rotations and users with the same orthogonal preamble collisions can be distinguished by the phase rotations. Note that the utilization of differential phase rotations not only expands the preamble pool, but also reduces the cross-correlation of multiple root ZC sequences, thus, reducing the probability of collisions and improving the performance of CE. At the BS, users without orthogonal preamble collisions can be identified by detecting the orthogonal sequences, and their channel vectors can be estimated accordingly. As for collided users, they are distinguished by processing non-orthogonal sequences to detect the phase rotations, and the channel vectors are estimated with low errors correspondingly.

In this paper, the scheme for constructing composite preambles by adding differential phase rotations to multiple root ZC sequences with the single cyclic shift is given first to illustrate the superiority of phase rotations, including the preamble structure, AUD and CE scheme, the theoretical performance of AUD and CE. Subsequently, the proposed scheme is extended to multiple root ZC sequences with randomly selected cyclic shifts, and the combination of phase rotations and cyclic shifts is given theoretically to further expand the preamble set size using cyclic shifts without increasing the non-orthogonal interference.

The novelty and contribution of this paper are summarized as follows.

1) The differential phase rotation-based composite preamble considering multiple root ZC sequences with the single cyclic shift is proposed, which reduces the non-orthogonal interference while enlarging the preamble set size by utilizing differential phase rotations.
2) Based on the proposed preamble, the detection algorithm and CE scheme are proposed, and the theoretical analysis of AUD and CE is derived to verify the superiority of differential phase rotations.
3) The extension to multiple root ZC sequences with randomly selected cyclic shifts is given, and the theoretical analysis of phase rotations-cyclic shifts combination is presented to further expand the preamble set size.

The rest of this paper is organized as follows. In Section II, the system model of GFRA system is illustrated. In Section III, the designed preambles and detection algorithm are proposed, and the performance analysis is also given. Section IV extends the proposed preamble to multiple root ZC sequences with randomly selected cyclic shifts. Simulation results are presented in section V. We conclude the paper in Section VI.

**Notation:** The matrices and vectors are denoted by boldface upper and lower case letters, respectively. $\mathbf{I}_M$ is the $M \times M$ unit matrix. $(\cdot)^T$ and $(\cdot)^H$ are the transpose and complex conjugate transpose operators. $\| \cdot \|$ denotes the Euclidean norm.
II. SYSTEM MODEL

Consider a single-cell grant-free access system as illustrated in Fig. 1, consisting of a BS equipped with $M$ antennas and multiple users. Each user is equipped with a single antenna, and it is assumed that there are $K$ active users uniformly distributed in the cell.

![Fig. 1. GFRA system model.](image)

As shown in Fig. 2, the user transmits preambles followed by the data to the BS, where the preambles are used for AUD and CE. The preambles are randomly selected from the preamble pool $\mathbf{S} = \{s_{l_1}, \ldots, s_{l_N}, \ldots, s_{l_L}\}^{N \times L}$ with $l_k = 1, \ldots, L$. Specifically, $s_{l_k}$ with index $l_k$ is the preamble selected by user $k$. $N$ is the length of preamble sequences and $L$ is the size of the preamble pool. If sequences $s_{l_k}$ ($l_k = 1, \ldots, L$) are orthogonal, $L = N$, otherwise, $L > N$. Generally, orthogonal preambles facilitate accurate user detection, but are prone to preamble collisions due to the limited size of the preamble pool. In contrast, non-orthogonal sequences can enlarge the preamble pool at the cost of decreased detection performance.

Denote $\mathbf{h}_k \in \mathbb{C}^{M \times 1}$ as the channel gain between user $k$ and the BS, where $\mathbf{h}_k = \sqrt{P_k} \mathbf{h}_{dk}$ with $\mathbf{h}_{dk}$ is the large-scale fading coefficient related to the distance and $\mathbf{h}_k \sim \mathcal{CN}(0, \mathbf{I}_{M \times 1})$ is the small-scale fading coefficient. Note that the assumption of uncorrelated fading has been widely used in GFRA systems for analysis and simulations [12], [14], [16]-[18]. In addition, it is assumed that the power control is perfect, which means that the received power $P = P_k |\mathbf{h}_{dk}|^2$ is identical for all active users even though their transmission power $P_k$ is different. Then, the received signal is

$$\mathbf{Y} = \sum_{k=1}^{K} \sqrt{P_k} \mathbf{h}_k s_{l_k} + \mathbf{W} \in \mathbb{C}^{M \times N},$$

where $\mathbf{W}$ is the the circularly symmetric complex Gaussian (CSCG) noise with zero-mean and variance $\sigma^2$. The signal-to-noise ratio (SNR) is $\gamma = \frac{P}{\sigma^2}$.

Fig. 2. GFRA transmission model.

Based on the received signal $\mathbf{Y}$, the BS performs AUD by detecting preambles selected by active users and also estimates channel coefficients to provide information for data detection.

III. PROPOSED PREAMBLE DESIGN

In this section, the composite preamble that assembles the advantages of orthogonal and non-orthogonal sequences is proposed to reduce the probability of preamble collision and thereby improve the performance of AUD and CE. The designed preambles are first given, followed by the detection scheme and performance analysis.

A. Designed preamble

The designed preamble is shown in Fig. 3, where $s_{l_k}$ consists of the orthogonal sequence $\mathbf{a}_{l_k}$ and non-orthogonal sequence $\mathbf{b}_{l_k, \theta_k}$. In detail, $\mathbf{a}_{l_k}$ is generated by the single-root ZC sequence, i.e.,

$$a_{l_k}(n) = \exp \left( -j \pi (n + l_k)(n + l_k + 1)/N_1 \right),$$

with $n = 1, \ldots, N_1$. $N_1$ is the length of $\mathbf{a}_{l_k}$ and $N_1 \geq N/2$. $l_k \in [0, N_1 - 1]$ is the cyclic shift. The single-root ZC sequence has the favorable property that $\|\mathbf{a}_{l_k} \mathbf{a}_{l_k}^H\|^2 = N_1$, but for $l_k' \neq l_k$, $\|\mathbf{a}_{l_k} \mathbf{a}_{l_k'}^H\|^2 = 0$.

As for the non-orthogonal sequence $\mathbf{b}_{l_k, \theta_k}$, it is generated by phase-rotating multiple root ZC sequences, which is expressed as

$$b_{l_k, \theta_k}(n) = \exp \left( -j \pi l_k (n + c_{l_k})(n + c_{l_k} + 1)/N_2 \right) - j n \theta_k,$$

with $n = 1, \ldots, N_2$. $N_2 = N - N_1$. To analyze the performance of phase rotations, it is first assumed that there is a one-to-one mapping relationship between $l_k$ and $c_{l_k}$. Parameter $\theta_k$ is the phase rotation, which is randomly selected from $\Theta = \{\theta_1, \theta_2, \ldots, \theta_Q\}$. Note that there is a relationship between $\mathbf{b}_{l_k, \theta_k}$ and $\mathbf{a}_{l_k}$, i.e., the cyclic shift of $\mathbf{a}_{l_k}$ is designated as the root of $\mathbf{b}_{l_k, \theta_k}$. Therefore, users selecting the same orthogonal preamble can be distinguished by the non-orthogonal preambles with different phase rotations. Phase rotation enlarges the size of the preamble pool from $N$ to $QN_1$, which reduces the probability of preamble collisions. Besides, it helps to reduce the cross-correlation of ZC sequences with different roots, which will be analyzed in detail in Section III-B.

With the composite preambles, the signal received by the BS is expressed as

$$\mathbf{Y} = [Y_1, Y_2],$$

Fig. 3. Designed composite preamble structure.
where
\[
Y_1 = \sum_{k=1}^{K} \sqrt{P} a_k^H a_k + W \in \mathbb{C}^{M \times N_1},
\]
(5)
and
\[
Y_2 = \sum_{k=1}^{K} \sqrt{P} b_k^H \theta_k + W \in \mathbb{C}^{M \times N_2}.
\]
(6)

### B. Preamble detection

After receiving the signal \( Y \) from the uplink transmission, the BS detects the preambles to identify the active users. First, BS processes \( Y_1 \) to detect which orthogonal sequences have been selected and initially determines whether they have been transmitted by multiple users. If the preamble has been selected by one user, the corresponding active user has been identified. Otherwise, it indicates that the preamble collision has occurred. These collided users are identified by detecting \( Y_2 \). The detailed detection scheme is as follows.

1) Orthogonal sequence detection: In view of the advantageous autocorrelation and cross-correlation properties of the single-root ZC sequence, \( a_k \) can be detected by calculating the correlation of \( Y_1 \) and \( a_k^H \), which is
\[
g_1 = Y_1^H a_k = \sum_{k=1}^{K} \sqrt{P} h_k a_k^H a_k^H + W a_k^H,
\]
(7)
If \( a_k \) has been selected,
\[
g_1 = \sum_{k \in L_k} \sqrt{P} h_k \|a_k\|^2 + W a_k^H,
\]
(8)
where \( L_k \) is the set of users which have selected \( a_k \) as the orthogonal preamble. Otherwise,
\[
g_1 = W a_k^H.
\]
(9)
The test statistic \( T_1 \) is defined as
\[
T_1 = \frac{1}{N_1} \|g_1\|^2.
\]
(10)
Denote the cases that \( a_k \) has not been selected and has been selected by only one user as \( H_0 \) and \( H_1 \), respectively. In the case of \( H_0 \), \( T_1 \) is a random variable that obeys a Chi-square distribution. According to the central limit theorem (CLT), \( T_1 \) can be approximated as a Gaussian distribution when \( N_1 \) is large [27]–[30], that is
\[
T_1 \sim \mathcal{N} (\lambda, \sigma^2).
\]
(11)
Similarly, under hypothesis \( H_1 \), \( T_1 \) is approximated by the following Gaussian distribution.
\[
T_1 \sim \mathcal{N} (MN_1 \sigma^2, 2MN_1 \sigma^2).
\]
(12)
Hence, by comparing \( T_1 \) with a threshold, the presence of \( a_k \) can be detected. The decision rule is given as
\[
\begin{align*}
H_0 : & \quad T_1 < \lambda_1 \\
H_1 : & \quad T_1 \geq \lambda_1
\end{align*}
\]
(13)
The threshold \( \lambda_1 \) satisfies
\[
\int_{T_1 \geq \lambda_1} f_{T_{1|H_1}} (\lambda_1) = \int_{T_1 < \lambda_1} f_{T_{1|H_0}} (\lambda_1).
\]
(14)
where \( f_{T_{1|H_0}} (\lambda_1) \) and \( f_{T_{1|H_1}} (\lambda_1) \) are the probability density function of \( T_1 \) in the case of \( H_0 \) and \( H_1 \), which are given as (11) and (37), respectively.

The probability of orthogonal sequence detection is derived as
\[
P_{d_1} = P_r (T_1 \geq \lambda_1 | H_1) = Q \left( \frac{\lambda_1 - MN_1 \sigma^2}{\sqrt{2MN_1 \sigma^2}} \right).
\]
(15)
If it is detected that \( a_k \) has been selected, then the next step is to determine if there exists the preamble collision. Denote the size of \( L_k \) as \( K_c \), which is calculated as
\[
K_c = \max \left( 1, \frac{T_1}{MN_1 \sigma^2} \right).
\]
(16)
If \( K_c = 1 \), \( a_k \) has been selected by only one user and the channel vector is estimated as
\[
\hat{h}_k = \frac{1}{N_1} Y_1 a_k^H.
\]
(17)
Otherwise, it is stated that there is a preamble collision.

Note that the only concern here is whether \( K_c \) is greater than 1, not the exact value of \( K_c \). This is because there may be false alarms, i.e., \( K_c \) is larger than the number of actually collided users. The false alarms ensure that the collided users will not be missed and they can be corrected in the non-orthogonal sequence detection.

2) Interference cancellation: Since users without preamble collisions have been detected and their channel vectors are estimated, the corresponding non-orthogonal preambles can be removed from \( Y_2 \) to cancel the interference caused by the non-collided users. The residual signal of \( Y_2 \) becomes
\[
Y_{2r} = Y_2 - \sum_{k \in L_c} \sqrt{P} h_k b_k^H \theta_k,
\]
(18)
where \( K_c \) is the index set of preambles chosen by non-collided users, and \( \theta_k \) is the estimated phase rotation selected by user \( k \). To estimate the phase rotation, the correlation between \( Y_2 \) and \( b_k^H \theta_k \) is considered, i.e.,
\[
Y_2 b_k^H \theta_k = \sqrt{P} h_k b_k^H \theta_k + \nu,
\]
(19)
where \( \nu \) is the interference caused by users who choose other orthogonal preambles and noise, and \( \nu \) is much smaller than \( |\sqrt{P} h_k b_k^H \theta_k| \) due to the property of multiple root ZC sequences. Therefore, \( Y_2 b_k^H \theta_k \) mainly depends on
\[
b_k^H \theta_k b_k^H \theta_k = \sum_{n=1}^{N_2} e^{j\pi(\theta - \theta_k) n},
\]
(20)
For the ease of representation, denote \( \Delta_\theta = \theta - \theta_k \). If and only if \( \Delta_\theta = 0 \), \( \sum_{n=1}^{N_2} e^{j\pi(\theta - \theta_k) n} = N_2 \), otherwise, it is negligible. For example, when \( \Delta_\theta = 0.5\pi \),
\[
\sum_{n=1}^{N_2} e^{j\pi(\theta - \theta_k) n} = \begin{cases} 0 & N_2 \text{ mod 4 } = 0 \\ j & N_2 \text{ mod 4 } = 1 \\ j - 1 & N_2 \text{ mod 4 } = 2 \\ -1 & N_2 \text{ mod 4 } = 3 \end{cases}
\]
(21)
where ‘mod’ is the modulo operation. Similar conclusions can be drawn when $\Delta_\theta$ is equal to other values. Therefore, the estimate of $\theta_k$ is obtained as

$$\hat{\theta}_k = \arg\max_{\theta \in \Theta} \| Y_2 b_k^H \|^2. \quad (22)$$

3) Non-orthogonal sequence detection: After canceling the interference of the non-collided users, the BS detects non-orthogonal preambles to identify the collided users based on $Y_2$. The set of collided preambles is represented as $K_n$, then $Y_2$ can be rewritten as

$$Y_2 = \sum_{l_k \in K_n} \sqrt{P_h} b_{l_k,\theta_k} + W. \quad (23)$$

As a result, the correlation of $Y_2$ with $b_{k,\theta_q}$ is

$$g_2 = Y_2 b_k^H = \sqrt{P_h} \| b_{l_k,\theta_q} \|^2 + \nu_{q_k} + \nu_{l_k} + W b_k^H; \quad (24)$$

where $\nu_{q_k}$ is the interference caused by users that choose the same $l_k$ but with different phase rotations, which can be expressed as

$$\nu_{q_k} = \sum_{q' \neq q} \sqrt{P_h} b_{l_k,\theta_{q'}} b_{l_k,\theta_{q'}}^H. \quad (25)$$

$\nu_{l_k}$ is the interference caused by the collided users choosing other orthogonal preambles, that is,

$$\nu_{l_k} = \sum_{l_k' \in K_n} \sqrt{P_h} b_{l_k',\theta_{l_k}} b_{l_k',\theta_{l_k}}^H. \quad (26)$$

As mentioned above, $|\nu_{q_k}|$ is negligible. While for $|\nu_{l_k}|$, it is closely related to $\Delta_\theta$, $\Delta_\nu = c_{l_k}' - c_{l_k}$ and $\Delta_{l_k} = l_k' - l_k$, which is much less than $N_2$. For example, when $N_2 = 32$, given $b_{1,0}(n) = \exp(-j \pi n (n + 1)/N_2)$, the cross-correlations of $b_{1,0}$ and sequences with various $l_k$, $c_{l_k}'$ and $\theta$ are shown in Table I. In detail, $\theta = 0$ indicates the case where there is no phase rotation and it is used as a benchmark. Values in red represent the case in which the cross-correlation is larger than the benchmark. It is shown that phase rotations reduce the cross-correlation in most cases. In particular, when $l_k' = 4$, $c_{l_k}' = 1$ and $l_k' = 2$, $c_{l_k}' = 2$, the cross-correlation is reduced by almost half. Even though at $l_k' = 2$, $c_{l_k}' = 1$ and $l_k' = 4$, $c_{l_k}' = 2$, $|\nu_{l_k}|$ at some phase rotations are larger than benchmarks, the average cross-correlation $|\nu_{l_k}|$ is close to or even less than the benchmark. Note that the conclusion also holds for situations with other values of $N_2$, $\Delta_\theta$, $\Delta_\nu$ and $\Delta_{l_k}$. Therefore, the correlation of non-orthogonal sequences can be decreased by phase-rotating the ZC sequences and thereby reducing the interference and improving the detection performance. Moreover, the preamble set size is enlarged by using phase rotations, which reduces the probability of preamble collision.

In the non-orthogonal sequence detection process, the test statistics is given as

$$T_2 = \frac{1}{N_2} \| g_2 \|^2. \quad (27)$$

Analogous to $T_1$, $T_2$ is approximated by the Gaussian distribution with mean $\mu_{T_2}$ and variance $\sigma_{T_2}^2$. Under the hypothesis $H_0$, i.e., $b_{k,\theta_q}$ has not been selected,

$$\mu_{T_2|H_0} = MP |\nu_{l_k}|^2 / N_2 + M \sigma^2$$

$$\sigma_{T_2|H_0}^2 = 2MP |\nu_{l_k}|^2 / N_2 + M \sigma^2. \quad (28)$$

Under the hypothesis $H_1$, i.e., $b_{k,\theta_q}$ has been selected,

$$\mu_{T_2|H_1} = MPN_2 + 2MP |\nu_{l_k}|^2 / N_2 + M \sigma^2$$

$$\sigma_{T_2|H_1}^2 = 2MPN_2 + 2MP |\nu_{l_k}|^2 / N_2 + M \sigma^2. \quad (29)$$

Similarly, the decision threshold $\lambda_2$ satisfies the condition that

$$f_{T_2|H_0}(\lambda_2) = f_{T_2|H_1}(\lambda_2). \quad (30)$$

The probability of non-orthogonal sequence detection is

$$P_{d_2} = Pr(T_2 \geq \lambda_2 | H_1)$$

$$= Q \left( \frac{\lambda_2 - MPN_2 - 2MP |\nu_{l_k}|^2 / N_2 - M \sigma^2}{\sqrt{2MPN_2 + 2MP |\nu_{l_k}|^2 / N_2 + M \sigma^2}} \right). \quad (31)$$

Since the phase rotations have been detected, the collided users can be identified and the channel vector can be estimated as

$$\hat{h}_k = \frac{1}{N_2 \sqrt{P} Y_2 s_{k,H}}. \quad (32)$$

with

$$s_{k} = [a_{l_k}, b_{l_k,\theta_k}]. \quad (33)$$

Note that for $l_k \in K_n$, if only one $\theta_k$ has been detected, it declares that $l_k \in K_n$, which indicates that there is a false alarm in the orthogonal sequence detection and it has been corrected in the non-orthogonal sequence detection. The whole detection process is summarized in Algorithm 1.

The complexity of orthogonal detection is $\mathcal{O}(MN_2^2)$. As for other steps, denote the size of $K_n$ as $K_n$ and $K_n$, where $K_n + K_n = K$. Then, the complexity of interference cancellation and non-orthogonal sequence detection are $\mathcal{O}(K_n MN_2 Q + MN)$ and $\mathcal{O}(K_n MN_2 Q)$, respectively. Therefore, the complexity of Algorithm 1 is $\mathcal{O}(MN_2^2 + KMN_2 Q + MN)$.
Algorithm 1 Detection Algorithm of The Preamble with Phase Rotation

Input: The received signal $Y = \{Y_1, Y_2\}$, decision threshold $\lambda_1, \lambda_2, K_o = 0, K_n = 0$.

Output: $K_o = \{k | a_k \text{ has been selected by one user}, \Lambda = \{(l_k, q_k) | b_k, \theta_k \text{ has been selected by the user colliding in choosing } a_k \}$.

Orthogonal sequence detection
1: for $l_k = 1 : L_1$ do
2: Calculate $T_1 = \frac{1}{N_1} \| Y_1 a_k^H \|^2$.
3: if $T_1 > \lambda_1$ then
4: $a_k$ has been used by $K_c = \max \left(1, \frac{T_1}{N_1} \right)$ users.
5: if $K_c = 1$ then
6: $l_k \in K_o$.
7: else
8: $l_k \in K_n$.
9: end if
10: else
11: $a_k$ has not been used.
12: end if
13: end for

Non-orthogonal sequence detection
14: while $l_k \in K_o$ do
15: Calculate $T_2 = \frac{1}{N_2} \| Y_2 b_k^H \theta_k \|^2$.
16: if $T_2 > \lambda_2$ then
17: $q_k$ has been used by $(l_k, q_k) \in \Lambda$.
18: else
19: $q_k$ has not been used.
20: end if
21: end while

Interference cancellation
22: while $l_k \in K_n$ do
23: for $q = 1 : \text{length } (\theta)$ do
24: $\theta_q = \Theta (q)$, $T_2 = \frac{1}{N_2} \| Y_2 b_k^H \theta_q \|^2$.
25: if $T_2 > \lambda_2$ then
26: $\theta_q$ has been used, $(l_k, q_k) \in \Lambda$.
27: else
28: $\theta_q$ has not been used.
29: end if
30: end for
31: end while

C. Performance analysis

With the proposed preambles, the probability that there is no preamble collision between $k$th user and other $K - 1$ users is

$$P_{nc} = \left(1 - \frac{1}{QN_1} \right)^{K-1}.$$

(34)

The probability that there is no orthogonal preamble collision between $k$th user and other $K - 1$ users is

$$P_{noc} = \left(1 - \frac{1}{N_1} \right)^{K-1}.$$

(35)

For the $k$th active user, the probability of detection is

$$P_{d_k} = P_{nc} P_{d_1} + (P_{nc} - P_{noc}) P_{d_2} P_{d_2},$$

where $P_{d_k}$ is the probability that $K_c > 1$, that is, $T_1 \geq \lambda_1$ and $T_1 \geq MN_1 P$. In other words, $T_1 \geq \lambda_{\text{max}}$, where $\lambda_{\text{max}} = \max (\lambda_1, MN_1 P)$. When user $k$ collides with $K_c - 1$ users, the mean and variance of $T_1$ are

$$\mu_{T_1} | K_c = K_c M N_1 P + M \sigma^2,$$

$$\sigma_{T_1}^2 | K_c = K_c (K_c - 1) M N_1^2 P + M \sigma^2.$$

Therefore, the probability that $K_c > 1$ is

$$P_{d_k} = Q \left( \frac{\lambda_{\text{max}} - K_c M N_1 P - M \sigma^2}{\sqrt{K_c (K_c - 1) M N_1^2 P + M \sigma^2}} \right).$$

(38)

The channel estimation error mainly results from the non-orthogonal sequences. In detail,

$$\hat{h}_k^n = \frac{1}{N_2^2} Y_2 b_k^H \theta_k = \hat{h}_k^n + \sum_{i=1}^{K_1} \nu_i \theta_i + \sum_{j=1}^{K_2} \nu_j + \frac{\widetilde{w}}{N_2^2},$$

(39)

$$= \hat{h}_k^n + \frac{K_1^2 \nu_i}{N_2} + \frac{K_2 \nu_j}{N_2} + \frac{\widetilde{w}}{N_2 \sqrt{P}}.$$

(40)

where $K_1$ is the number of users that select preambles with the same $l_k$. $K_2$ is the number of users that also suffer from preamble collisions despite choosing other orthogonal preambles. $\nu_i$ and $\nu_j$ are the mean of $\nu_\theta_i$ and $\nu_\theta_j$, respectively. $\widetilde{w} = \text{Wb}_k^H \theta_k \sim \mathcal{CN}(0, I_M)$.

The channel estimation error is calculated as

$$e_h = \hat{h}_k^n - h_k^n = \frac{K_1^2 \nu_i}{N_2} + \frac{K_2 \nu_j}{N_2} + \frac{\widetilde{w}}{N_2 \sqrt{P}}.$$

(40)

As mentioned in Section III-B, $|\nu \theta_i|$ is generally reduced by the differential phase rotations. In other words, the CE errors caused by non-orthogonal sequences depending on $\nu_\theta_i$ and $\nu_\theta_j$ are reduced.

IV. EXTENSION OF THE PROPOSED PREAMBLE

The above analysis assumes adding phase rotation to the multiple root ZC sequences with a zero cyclic offset, but the scheme can be also extended to multiple root ZC sequences with different cyclic shifts. Then the non-orthogonal preamble is given as

$$b_{l_k, c_k, \theta_k} (n) = \exp \left( -j \pi l_k (n + c_k) (n + c_k + 1) N_2 \right) \frac{\theta_k n }{N_2},$$

(41)

where $c_k$ is the cyclic shift randomly selected in the range of $[1, C_{\text{max}}]$, and $\theta_k$ is randomly selected from the set $\Theta$.

It is obvious that the cross-correlation $|b_{l_k, c_k, \theta_k} b_{l_k, c_k, \theta_k}^H|$ is related to $l_k, c_k, \theta_k - \theta_k$. To avoid the worst situation that $|b_{l_k, c_k, \theta_k} b_{l_k, c_k, \theta_k}^H| = N_2$, with given $l_k$, the selection of the cyclic shift and phase rotation should satisfy Lemma 1.
Lemma 1: For any $\theta_q, \theta_k \in \Theta$ and any $c_q, c_k \in [1, C_{\text{max}}]$, the following condition should to be met.

$$
2l_k \pi \left( c_q - c_k \right) + \theta_q - \theta_k \neq 2\alpha \pi, \alpha = 0, 1, 2, \ldots
$$

Proof: Due to the properties of ZC sequences, sequences with different roots can be detected by comparing the correlation with a threshold. But for sequences with the same root but different $c_k$ or $\theta_k$, the cross-correlation is calculated as

$$
G = \left| b_{l_k, c_q, \theta_q}^H b_{\theta_k, c_k, \theta_k} \right|
$$

$$
= \sum_{n=1}^{N_2} \exp \left( jn \left( \frac{2l_k \pi \Delta_c}{N_2} + \Delta_\theta \right) + \frac{jl_k \pi D}{N_2} \right)
$$

Here, $\Delta_c = c_q - c_k$, $\Delta_\theta = \theta_q - \theta_k$ and $D = \Delta_c \left( c_q + c_k + 1 \right)$. $G$ is much less than $N_2$ in the case of $c_q = c_k, \theta_q \neq \theta_k$ and $c_q \neq c_k, \theta_q = \theta_k$. However, when $\frac{2l_k \pi \Delta_c}{N_2} + \Delta_\theta = 0, 2\pi, 4\pi, \ldots$, $G = N_2$ even if $c_q \neq c_k, \theta_q \neq \theta_k$, which leads to false detections. Therefore, to avoid this situation, $\Delta_c$ and $\Delta_\theta$ are supposed to satisfy that

$$
\frac{2l_k \pi \Delta_c}{N_2} + \Delta_\theta \neq 2\alpha \pi, \alpha = 0, 1, 2, \ldots
$$

For example, assuming that $\Theta = [0, \pi/2, \pi, 3\pi/2]$ and $l_k = 1$, (42) is

$$
\Delta_c \neq \alpha N_2 - \frac{\Delta_\theta N_2}{2\pi}, \alpha = 0, 1, 2, \ldots
$$

Since $\Delta_c < N_2$ and $\Delta_\theta < 2\pi$, (45) always holds when $\alpha > 1$. Then the problem is to find $C_{\text{max}}$ so that $\Delta_c \neq -\Delta_\theta N_2 / 2\pi$ and $\Delta_c \neq N_2 - \Delta_\theta N_2 / 2\pi$. With the given $\Theta$,

$$
\left| \frac{\Delta_\theta}{2\pi} \right| = \left\{ \frac{N_2 - 3N_2}{4}, \frac{N_2 - 7N_2}{4}, \frac{N_2 - 3N_2}{4} \right\},
$$

$$
\left| \frac{\Delta_\theta}{2\pi} \right| = \left\{ \frac{N_2 - 3N_2}{4}, \frac{N_2 - 7N_2}{4}, \frac{N_2 - 3N_2}{4} \right\}.
$$

Therefore, it is derived that $\Delta_c < \frac{N_2}{4}$ and thus $C_{\text{max}} = \left\lceil \frac{N_2}{4} + 1 \right\rceil$.

Similarly, the setting of $\Theta$ and $C_{\text{max}}$ for other values of $l_k$ can be derived.

V. SIMULATION RESULTS

In this section, simulations are performed to evaluate the performance of the designed preamble. Simulation results in terms of the probability of AUD, normalized mean square error (NMSE) of CE and success rate defined in [26] are presented to verify the effectiveness of the proposed composite preamble. The orthogonal preambles, auxiliary preambles assisted by multiple root ZC sequences without phase rotations [15] and aggregate preambles based on orthogonal sequences [22] are chosen as comparisons.

A. Performance of AUD

The probability of no preamble collision is given in Fig. 4, where $N$ is fixed as 64. It is shown that the probability of no preamble collisions $P_{nc}$ is much increased with differential rotations, especially when the number of active users is larger than the preamble length, that is $K > 60$. In particular, the more phase rotations available and the longer the orthogonal sequence, i.e., the larger the $Q$ and $N_1$, the larger $P_{nc}$. However, although collisions can be reduced by increasing the length of orthogonal sequences, the performance of AUD cannot be consistently improved as $N_1$ increases.

The variation of detection performance with $N_1$ is shown in Fig. 5, where the numerical results are consistent with the theoretical ones. The length of the preamble is $N = 64$, and phase rotations are randomly selected from $\Theta = \left\{ \pi/4, \pi/2, 3\pi/4, \pi, 5\pi/4, 3\pi/2, 7\pi/4, 2\pi \right\}$. It is obvious that there is an optimal $N_1$ that maximizes $P_d$, and the optimal $N_1$ is closely related to communication conditions. First, the performance is improved with increasing $N_1$, because the longer orthogonal sequence enlarges the preamble set size and contributes to accurate detection. However, the length of non-orthogonal sequences $N_2$ is decreased with $N_1$, which limits non-orthogonal sequence detection accuracy. In addition, as $N_1$ increases, $N_1$ has a diminishing effect on reducing the probability of orthogonal preamble collisions, and the performance of AUD mainly depends on non-orthogonal sequences. Therefore, the performance first increases with $N_1$, but when $N_1$ has little effect on reducing orthogonal preamble collisions and it reduces the non-orthogonal sequence detection performance significantly, the detection accuracy decreases as the orthogonal sequences become longer.
to some extent but also reduces the performance orthogonal sequences. As for $\Theta$, it is used to enlarge the preamble pool at the cost of introducing more nonorthogonality, and the smaller $\Delta_\Theta$, the larger of the preamble pool. The importance of orthogonal or non-orthogonal detection depends on the number of users. When $K$ is small, orthogonal detection plays a dominant role, and smaller $\Delta_\Theta$ makes no difference since it is unnecessary to further expand the preamble pool. However, when $K$ is as large as 60, orthogonal preamble collisions are inevitable and $\Theta$ is needed. It is worth noting that when $N_2 \leq 20$, larger $\Delta_\Theta$ performs better, while things are quite different when $N_2 > 20$. That is because in the case of $N_2 \leq 20$, i.e., $N_1 \geq 44$, the probability of preamble collision is low and users can be detected as long as users colliding in orthogonal preambles can be identified, which indicates that improving the detection performance of non-orthogonal detection is more important than expanding the preamble pool. Since the overall interference of $\Delta_\Theta = \frac{\pi}{4}$ is smaller, it performs better. But when $N_2 > 20$, the probability of preamble collision is high, and solving the collision problem is more critical. Therefore, $\Delta_\Theta = \frac{\pi}{4}$ outperforms $\Delta_\Theta = \frac{\pi}{2}$ even though its interference is larger.

Fig. 7 compares the performance of AUD of the proposed preamble with other schemes, where $M = 10$, $\gamma = -5$ dB. The preamble length and phase rotations set are $N_1 = N_2 = 32$ and $\Theta = [\pi/4, \pi/2, 3\pi/4, \pi, 5\pi/4, 3\pi/2, 7\pi/4, 2\pi]$. Without avoiding preamble collisions, the performance of preambles used in LTE decrease dramatically due to the high probability of collisions, which is not the case with other schemes that consider preamble collisions. As for preambles considering collisions avoidance, the proposed preamble performs better than the other two, where the numerical results accord with the theoretical ones. In particular, the performance advantage of composite preambles over auxiliary preambles becomes more significant as the number of active users increases. That is because the number of collided users increases with $K$, and so does the non-orthogonal interference. In such cases, with the help of differential phase rotations, the non-orthogonality is reduced in the designed preamble, which cannot be mitigated in secondary preambles without phase rotations.

**B. Performance of CE and success rate**

The performance of CE is measured by the average NMSE, which is calculated as

$$\text{NMSE} = \frac{1}{I} \sum_{i=1}^{I} \frac{\|\hat{h}_{k,i} - h_{k,i}\|^2}{\|h_{k,i}\|^2},$$  \hspace{1cm} (47)

where $\hat{h}_{k,i}$ and $h_{k,i}$ are the estimated and exact channel vectors of user $k$ in $i$th Monte Carlo simulation, respectively. $I$ is the number of Monte Carlo simulations. The simulation results under different SNR conditions are presented in Fig. 8, where $M = 40$, $K = 50$, and $N_1 = 40$, $N = 64$. $\Theta$ is the same as Fig. 7. It’s shown that the estimation errors of the proposed preamble is less than other preambles. It is worth noting that the NNSE curves of the proposed preamble and aggregated preamble are getting closer and closer when SNR > 0 dB, and they almost intersect at SNR = 6 dB.
Since two orthogonal preambles with different power levels are adopted in the aggregated preamble, the detection performance of preambles under low SNR conditions cannot be guaranteed due to the inaccurate detection of power levels, which leads to the larger deviation of CE. However, as the SNR increases, the detection probability of aggregated preambles gradually approaches 1. In addition, as a consequence of random combinations of orthogonal sequences and power levels, its preamble set size is greatly enlarged and the probability of preamble collisions is the lowest among the four mentioned schemes. Therefore, the CE error of aggregated preambles is pretty low under high SNR conditions. As for the proposed composite preambles, even though the probability of no collisions is lower than aggregated preambles, it is still close to 0.9, which is sufficient for a low estimation error with the assistance of higher detection probability. Hence, the NMSE of proposed composite preamble is much lower than benchmarks under low SNR conditions, and it gradually approaches the extremely low level that is close to the NMSE of aggregated preambles under high SNR conditions.

![Fig. 8. NNSE of CE under various SNR conditions.](image)

The NMSE of CE vs. the number of active users is plotted in Fig. 9, where SNR = −5 dB and other simulation parameters are set to be the same as Fig. 8. It is obvious that the more active users, the larger NNSE, which is caused by the increased probability of preamble collision. In addition, in terms of the increasing rate of NNSE with respect to \( K \), the proposed preamble outperforms the auxiliary and orthogonal preambles due to the more enlarged preamble set size. Furthermore, compared with the aggregated preambles, the proposed preamble is slightly worse in increasing rate since its rate is larger, but it performs better owing to the higher detection probability.

The success rate of the proposed preamble and comparisons under conditions with different SINR thresholds are shown in Fig. 10, where \( \mathrm{SNR} = 0 \) dB and other parameters are set to be the same as the above simulations. The success rate is defined in [26] to evaluate the success access probability, which is mainly related to the probability of preamble collision and CE performance. Specifically, the success rate of \( k_{th} \) user is calculated as

\[
P_{\text{suc}} = P_{\text{nc}} P \left( \gamma_k^1 > \gamma_{th}^1 \right),
\]

where \( \gamma_k^1 \) is the received signal to interference and noise ratio (SINR), which is calculated based on the recovered signal

\[
y_k = h_k^H Y.
\]

\( \gamma_{th}^1 \) is the SINR threshold. Similarly, since the proposed composite preamble reduces preamble collision with low non-orthogonality, it is superior to other preambles in terms of success rate.

C. Performance of the proposed preamble with multiple cyclic shifts

The performance of the composite preambles with multiple cyclic shifts and phase rotations in terms of the probability of detection and success rate are given in Fig. 11 and Fig. 12, respectively. Here, \( M = 40 \), \( \mathrm{SNR} = 0 \) dB and the preamble length is \( N = 64 \) with \( N_1 = 32 \). Phase rotations are randomly selected from \( \Theta_1 = [0, \pi/4, \pi/2, 3\pi/4, \pi, 5\pi/4, 3\pi/2, 7\pi/4] \) or \( \Theta_2 = [0, \pi/2, \pi, 3\pi/2] \), and the cyclic shifts are obtained based on Lemma 1 accordingly. It’s shown that the combination of cyclic shifts and phase rotations improves the probability of detection and success rate. Note that the performance improvement is more significant in the case of \( \Theta_2 \), which can be explained by the fact that \( C_{\text{max}} \) is related to \( \Delta \), and the larger the \( \Delta \), the more cyclic shifts are available.

![Fig. 9. NNSE of CE with different number of active users.](image)

VI. CONCLUSION

In this paper, a composite preamble consisting of orthogonal sequences and non-orthogonal sequences with differential phase rotations is designed to reduce the non-orthogonal interference while reducing the probability of preamble collisions, and thereby improving the performance of AUD and CE. Accordingly, the detection algorithm and CE scheme are also proposed, and the performance is analyzed theoretically. It is shown that differential phase rotations enables the expansion...
of preamble set size and the reduction of non-orthogonality. Simulation results show that the designed preamble performs better than orthogonal sequence-based preambles and non-orthogonal preambles without phase rotations.

For the design and extension of the composite preamble, there are some other issues to be studied. For example, as shown in Fig. 5 and Fig. 6, there always exists an optimal length of orthogonal sequences that maximizes the detection performance under different conditions, and the optimal length has an influence on the design of the phase set. Therefore, further research is needed to adaptively determine the optimal length of orthogonal sequences and phase sets according to the conditions. In addition, the preamble is proposed for the single-cell system, whose performance degrades in the multi-cell scenario. To extend the proposed scheme to multi-cell scenarios, the design of the composite preamble needs to be improved to reduce the inter-cell interference, and introducing cell-specific power coefficients in the proposed preamble is a promising way. Moreover, GFRA has great potential in other scenarios, including ultra-reliable low-latency communications (URLLC), cell-free MIMO, and unmanned aerial vehicle (UAV) networks, etc. How to solve the preamble transmission and detection in these scenarios is still an open issue.

REFERENCES


Jiaru Lin (Member, IEEE) received the B.S. and Ph.D. degrees from the School of Information Engineering, Beijing University of Posts and Telecommunications (BUPT), China, in 1987 and 2001, respectively. From 1991 to 1994, he studied at the Swiss Federal Institute of Technology, Zürich. He is currently a Professor and the Ph.D. Supervisor with the School of Artificial Intelligence, BUPT. His research interests include wireless communications, personal communication, and communication networks.

Miao Pan (S07-M12-SM18) received his BSc degree in Electrical Engineering from Dalian University of Technology, China, in 2004, MASc degree in electrical and computer engineering from Beijing University of Posts and Telecommunications, China, in 2007 and Ph.D. degree in Electrical and Computer Engineering from the University of Florida in 2012, respectively. He is now an Associate Professor in the Department of Electrical and Computer Engineering at University of Houston. He was a recipient of NSF CAREER Award in 2014. His research interests include Wireless/AI for AI/Wireless, deep learning privacy, cybersecurity, and underwater communications and networking. His work won IEEE TCGCC (Technical Committee on Green Communications and Computing) Best Conference Paper Awards 2019, and Best Paper Awards in ICC 2019, VTC 2018, Globecom 2017 and Globecom 2015, respectively.

Dr. Pan is an Editor for IEEE Open Journal of Vehicular Technology, an Associate Editor for ACM Computing Surveys and an Associate Editor for IEEE Internet of Things (IoT) Journal (Area 5: Artificial Intelligence for IoT), and used to be an Associate Editor for IEEE Internet of Things (IoT) Journal (Area 4: Services, Applications, and Other Topics for IoT) from 2015 to 2018. He has also been serving as a Technical Organizing Committee for several conferences such as TPC Co-Chair for Mobiquitous 2019, ACM WUWNet 2019. He is a member of AAAI, a member of ACM, and a senior member of IEEE.