

# Low-Correlation Superimposed Pilot Grant-Free Massive Access for Satellite Internet of Things

Liang Xu, Jian Jiao<sup>ID</sup>, *Senior Member, IEEE*, Ye Wang<sup>ID</sup>, *Member, IEEE*, Shaohua Wu<sup>ID</sup>, *Member, IEEE*, Rongxing Lu<sup>ID</sup>, *Fellow, IEEE*, and Qinyu Zhang<sup>ID</sup>, *Senior Member, IEEE*

**Abstract**—Satellite Internet of Things (S-IoT) with low Earth orbit satellites has become an effective solution for providing global coverage for massive machine type communications (mMTC). Considering that the massive user equipments covered by the S-IoT are periodically activated and dominated by short packet communications, the pilot collision has become a challenging problem due to the limited length and number of pilot sequences. In this paper, we propose a low-correlation superimposed pilot grant-free massive access (LSP-GFMA) scheme, where a low-correlation-zone periodic sequence (LPS) is designed for the superimposed pilot (SP) structure. Our LPS can maintain low cross-correlation with random non-orthogonal shifts compared with the conventional Zadoff-Chu sequence (ZCS), which can alleviate pilot collision while ensuring high spectral efficiency. In addition, we propose an iterative channel estimation based on Kaczmarz algorithm to attain accurate channel state information for the SP structure with low complexity. Then, we derive the theoretical expressions of access failure probability (AFP) and achievable throughput for our LSP-GFMA scheme under the shadowed-Rician fading channel. Simulation results validate the accuracy of our theoretical derivations, and demonstrate that our LSP-GFMA scheme with LPS can achieve lower AFP and higher achievable throughput than that with ZCS, and also outperforms the state-of-the-art schemes.

**Index Terms**—Satellite Internet of Things, grant-free random access, low-correlation-zone periodic sequence, superimposed pilot, iterative channel estimation.

Manuscript received 20 February 2023; revised 6 June 2023 and 10 August 2023; accepted 4 September 2023. Date of publication 11 September 2023; date of current version 19 December 2023. This work was supported in part by the National Natural Sciences Foundation of China (NSFC) under Grant 62071141, Grant 61831008, and Grant 62027802, in part by the Shenzhen Science and Technology Program under Grant JSGG20220831110801003, and in part by the Major Key Project of PCL Department of Broadband Communication. The associate editor coordinating the review of this article and approving it for publication was H. Tabassum. (*Corresponding authors: Jian Jiao; Qinyu Zhang.*)

Liang Xu, Jian Jiao, Shaohua Wu, and Qinyu Zhang are with the Guangdong Provincial Key Laboratory of Aerospace Communication and Networking Technology, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China, and also with the Peng Cheng Laboratory, Shenzhen 518055, China (e-mail: xuliang\_hit@foxmail.com; jiaojian@hit.edu.cn; hitwush@hit.edu.cn; zqy@hit.edu.cn).

Ye Wang is with the Peng Cheng Laboratory, Shenzhen 518055, China (e-mail: wangy02@pcl.ac.cn).

Rongxing Lu is with the Faculty of Computer Science, University of New Brunswick, Fredericton, NB E3B 5A3, Canada (e-mail: rlu1@unb.ca).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TCOMM.2023.3313614>.

Digital Object Identifier 10.1109/TCOMM.2023.3313614

## I. INTRODUCTION

WITH the rapid development of low Earth orbit (LEO) satellite constellation, the satellite Internet of Things (S-IoT) is regarded as a promising solution to enable ubiquitous connectivity in global coverage [1]. The S-IoT can support the massive machine-type communications (mMTC) in the upcoming fifth generation-advanced (5G-A) network, such as intelligent transport systems, aeronautical telecommunication, international multimodal transport, and environmental monitoring, etc [2], [3]. However, the massive connectivity for uplink S-IoT and large propagation delay in satellite-ground links make the massive access in S-IoT become a challenging issue [4], [5].

Starting from the initial publication of ALOHA, a vast amount of evolutions of random access protocols can be found in the literatures [6], [7]. The Contention Resolution Diversity Slotted Aloha (CRDSA) protocol is proposed in [6], where the terminal is allowed to generate two replicas of the same packet. In the CRDSA scheme, the recovered information from a successful packet is exploited to cancel the interference that its twin may generate on another time slot. This approach is iterated to recover most of the frame packets that were initially lost due to collision. The enhanced spread spectrum Aloha (E-SSA) scheme is another approach to enhance the performance of packet-based spread-spectrum ALOHA (SSA) random access [7]. The E-SSA employs successive interference cancelation (SIC) at the packet level with a cyclic redundancy check (CRC) and opportunistic transmission control mechanism (OTCM).

Recently, the grant-free random access scheme is considered for massive access in S-IoT because of reducing handshake phases and control overhead [8], [9]. As the traffic of random access is sporadic, the satellite access point (SAP) may have no prior information on the activity of the user equipment (UE) nor the channel state information in each random access slot. In general, the UE in GFMA randomly selects a pilot sequence prior to data and directly transmits the pilot to the SAP along with its data, and the SAP differentiates UEs by the pilot sequences. However, a conventional collision occurs when more than one UEs select the same pilot in GFMA and would deteriorate the UE detection [10], [11]. Many research works focus on pilot signals design to mitigate this problem. The Gold codes constructed by a pair of m-sequences, have unique autocorrelation properties and are widely applied in CDMA

systems. The Zadoff-Chu (ZC) sequences with different cyclic shifts and root indexes are utilized for pilot design in the 3GPP LTE system [12]. To alleviate pilot collision, a non-orthogonal spreading sequence is considered a complement to pilot resources [13], [14]. However, although introducing the non-orthogonal spreading sequence can decrease pilot collision probability, it causes interference between non-orthogonal UEs, which has negative effect on channel estimation and access failure probability (AFP) [15]. Thus, the design of low correlation pilot sequence is a key part of GFMA scheme.

Moreover, the traffic in GFMA is characterized by short packet communications, where the regular pilot (RP) structure (i. e., the pilot sequence and data are successively transmitted) may deteriorate the spectral efficiency [16]. In the superimposed pilot (SP) structure, the pilot sequences are superimposed on data sequences and transmit together, which do not take up additional time-frequency resources [17]. However, we need to design a low complexity channel estimation algorithm to attain perfect channel state information (CSI) due to the inherent data interference in SP structure.

#### A. Related Works

The largest number of available orthogonal pilot sequences with length  $L$  is equal to  $L$  [18]. However, the orthogonal pilot sequence set cannot fully alleviate the pilot collision in GFMA. To enlarge the pilot sequence set, the authors in [19] exploit the potential of ZCS with its non-orthogonal sequences from at most  $(L - 1)$  roots. Besides, the  $m$ -sequence and all-top sequence also can be employed to multi-root ZCS (mZCS) [21], [22], which increases the number of pilots by  $(L + 1)$  times. Note that each pair of non-orthogonal ZCS have the same cross-correlation  $1/\sqrt{L}$ , which leads to serious interference between non-orthogonal UEs, i.e., the UEs who select the non-orthogonal ZCS. In [23] and [24], the non-orthogonal interference can be relieved in the receiver, which is designed through modeling sparse signal recovery into a multiple-measurement vector class of compressing sensing problem. To further reduce the non-orthogonal interference, a secondary pilot structure is proposed in [25], where the orthogonal sequences in mZCS set are used as the main pilot for UE detection, and the non-orthogonal sequences in mZCS set are used for alleviating main pilot collision as a secondary pilot. Thus, a UE is only disturbed by the interference from other UEs who select the same orthogonal pilot. However, the secondary pilot structure may deteriorate the spectral efficiency in short packet communication scenario.

The SP structure has the potential to improve the spectral efficiency in short packet communications for mMTC [26]. On one hand, the existing channel estimation algorithms for SP structure are mainly based on the grant-based random access protocol [27], [28], [29]. On the other hand, the channel estimation algorithms for SP structure in [30], [31] and [32] can be utilized in grant-free random access. In [30] and [31], the non-orthogonal interference can be significantly alleviated by using the subspace projection combined with the independent component analysis. Besides, the authors in [32] propose an estimator by strictly following the definition

of the minimum mean square error (MMSE) estimation, which eliminates the inherent data interference in SP structure significantly and achieves a similar MMSE. However, these methods in [30], [31] and [32] depend on the MMSE decoder, and the corresponding complexity is  $O(K^3)$  due to the large dimensional matrix inversion for each UE, where  $K$  denotes the number of activated UEs.

Due to the large propagation delay in high dynamic LEO satellite-ground links and limited onboard resources, an effective channel estimation algorithm is required for the GFMA in S-IoT. An iterative channel estimation based on ridge regression (ICER) algorithm is proposed in [33], which can achieve fast convergence by several iterations. However, the ICER algorithm needs the inversion operation for the reconstruction of signal matrix, which still leads to  $O(K^3)$  computation complexity. Moreover, the inevitable channel estimation error in SP structure would deteriorate the performance of successive interference cancellation (SIC) for UE detection [34], [35]. Then, the effect of imperfect SIC for the GFMA scheme in S-IoT should be further analyzed.

#### B. Contributions

In this paper, we propose a low-correlation superimposed pilot GFMA (LSP-GFMA) scheme for the uplink mMTC in S-IoT, and design a joint iterative channel estimation based on Kaczmarz (ICEK) and SIC decoding algorithm. The main contributions of the paper are summarized as follows.

- By utilizing the Hadamard matrix of order  $M = M' \times N'$  with elements  $\{-1, 1\}$  and a perfect sequence with length  $N$  with elements  $\{-1, 0, 1\}$ , we can generate a  $M \times N' \times N$  low-correlation-zone periodic sequence (LPS) set with  $M$  root LPS (rLPS) of length  $L = M \times N$ , and each rLPS can generate  $N' \times N$  shifted LPS (sLPS) by left-shifting each bit of the first  $N' \times N - 1$  bits in rLPS. Thus, we utilize rLPS set as the pilot set in our LSP-GFMA scheme, where each activated UE randomly selects and shifts a rLPS, and then performs random access to the SAP, which can significantly decrease the pilot collision probability. Further, we prove that the average auto-correlation and cross-correlation of our LPS is  $O(1/L)$ , which is  $1/\sqrt{L}$  times lower than that of mZCS, and can improve the decoding failure probability over the shadowed-Rician fading channel.
- Our LSP-GFMA scheme adopts the SP structure with LPS set to improve the spectral efficiency in short packet communications. In order to reduce the inevitable channel estimation error caused by SP structure, we firstly propose an ICEK algorithm, which can attain the similar performance of normalized mean square error (NMSE) as the conventional MMSE estimation with complexity  $O(K^2)$ . Then, we propose a joint ICEK and SIC decoding algorithm to further improve the accuracy of CSI and decoding failure probability, where the intermediate decoded data from the SIC decoder is fed back to the ICEK to update the estimated CSI iteratively.
- We derive the probability density function (PDF) of the interference from non-orthogonal UEs over the

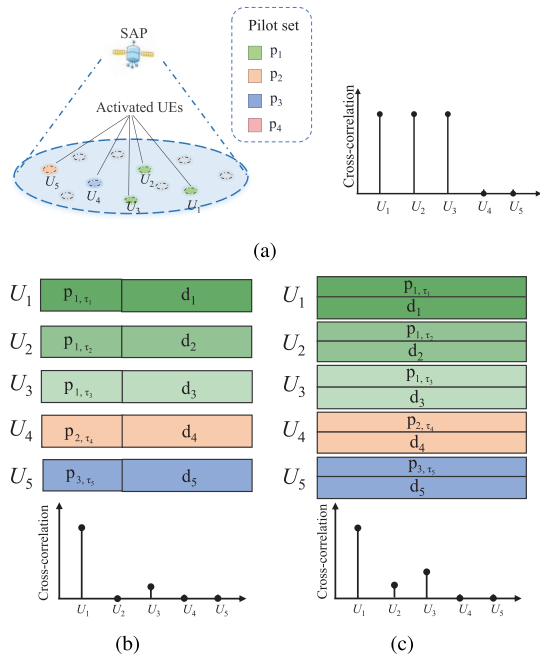


Fig. 1. The uplink GFMA S-IoT scenario, (a) conventional orthogonal pilot set, (b) LPS pilot set with random shifting in RP structure, (c) the LSP-GFMA scheme.

shadowed-Rician fading channel, and further derive the expression of signal to interference noise ratio (SINR) by taking into account the imperfect SIC at SAP and the channel estimation error after the joint ICEK and SIC decoding algorithm. Thus, we have the closed-form expression of the decoding failure probability. Finally, we derive the closed-form expression of the AFP in LSP-GFMA scheme, which is mainly determined by the practical decoding failure probability, and further derive the achievable throughput from the AFP. Simulations validate the accuracy of the theoretical derivations, and demonstrate that our LSP-GFMA scheme has superior AFP and achievable throughput performances in short packet communication.

C. Paper Organization

The rest of this paper is organized as follows. Section II introduces the system model and LSP-GFMA scheme. Section III describes the ICEK algorithm for SP structure, and details the design methods and characteristics of LPS. In section IV, we derive the theoretical expressions of AFP and achievable throughput over the shadowed-Rician fading channel. Section V shows simulation results, and conclusions are drawn in Section VI.

II. SYSTEM MODEL AND LSP-GFMA SCHEME

In this section, we first outline the system model and analyze the pilot collision. Then, we propose the LSP-GFMA scheme to alleviate the collision.

A. System Model

The uplink GFMA S-IoT scenario is depicted in Fig. 1 (a), where a Ka band SAP with an altitude of 600 km provides

mMTC services for UEs of a spot beam coverage [3], [4], [36]. The UEs are quasi-stationary and uniformly random distributed in the spot beam coverage area. Since the satellite-ground link is larger than several hundred kilometers away, we assume that each UE has approximately the same distance from SAP and arrives at almost the same time. One random access time slot consists of the propagation delay and processing delay at the transceiver. Since the worst two-way propagation delay is expected to be 26 ms for the SAP at 600 km [37], the random access time slot is set as 50 ms [38]. In each random access time slot, the covered UEs are activated with activation probability  $p_a$  to perform access [39], and  $K$  activated UEs randomly select the pilot from a size  $M$  orthogonal pilot set, and transmit data frame in length  $B$  to the SAP. Assume the pilot length is  $L$ , then the pilot set can be represented as  $\mathbf{P} = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_M\} \in \mathbb{C}^{L \times M}$ . The pilot collision will occur when multiple UEs select the same pilot. For example, three UEs  $U_1, U_2$  and  $U_3$  randomly select the same pilot  $\mathbf{p}_1$  as shown in Fig. 1, where the y-axis in Fig. 1 corresponds to the cross-correlation of a pair of pilots that are selected by UE  $U_1$  and the other UEs.

To alleviate the pilot collision, we construct a novel LPS pilot set with  $M$  orthogonal rLPS for this uplink GFMA S-IoT scenario, where both average auto-correlation and cross-correlation of our LPS are  $O(1/L)$ , which can significantly decrease the pilot collision probability, and also achieves lower decoding failure probability than the mZCS over the shadowed-Rician fading channel. The construction of LPS is detailed in Section III-B.

In our LSP-GFMA scheme, each activated UE randomly selects and shifts a rLPS to obtain a sLPS as shown in Fig. 1(b) and Fig. 1(c), where  $U_1, U_2$  and  $U_3$  randomly choose the same rLPS  $\mathbf{p}_1$  and left shift  $\tau_1, \tau_2$  and  $\tau_3$  symbols on  $\mathbf{p}_1$  to obtain  $\mathbf{p}_{1,\tau_1}, \mathbf{p}_{1,\tau_2}$  and  $\mathbf{p}_{1,\tau_3}$ , respectively. Then, these sLPS pilots are transmitted with the data  $\mathbf{d}_1, \mathbf{d}_2$ , and  $\mathbf{d}_3$  to the SAP. Thanks to the low average cross-correlation of pilot in sLPS set, the SAP can distinguish  $U_1, U_2$  and  $U_3$ , and then recover the data message.

Moreover, consider the dominant traffic in uplink GFMA is short packet communication, the conventional RP structure in Fig. 1(b) may lead to low spectral efficiency when the ratio of  $B$  to  $L$  is low [32]. Thus, we propose a SP structure as shown in Fig. 1(c), where the sLPS pilots  $\mathbf{p}_{1,\tau_1}, \mathbf{p}_{1,\tau_2}$  and  $\mathbf{p}_{1,\tau_3}$  are superimposed on the data  $\mathbf{d}_1, \mathbf{d}_2$ , and  $\mathbf{d}_3$ , respectively, and transmitted to the SAP. The SAP can obtain UE's identification information by signature information embedded in data header [40]. Further, we propose an ICEK algorithm to deal with the inevitable channel estimation error caused by SP structure, which is detailed in Section III-A. Then, the received superimposed signal at SAP can be expressed as

$$\mathbf{y} = \sum_{i=1}^K g_i(\sqrt{\alpha P_t} \mathbf{d}_i + \sqrt{(1-\alpha)P_t} \mathbf{p}_{i,\tau_i}) + \omega, \quad (1)$$

where  $\mathbf{d}_i$  and  $\mathbf{p}_{i,\tau_i}$  are data and a sLPS that has  $\tau_i$  symbols left shift on  $\mathbf{p}_i$  of  $U_i$ , respectively,  $P_t$  represents total transmit power,  $\alpha$  represents power allocation coefficient,  $\omega$  denotes the additive white Gaussian noise (AWGN) as  $\mathcal{CN}(0, \sigma^2)$ . Further,

$g_i$  represents the channel gain coefficient between  $U_i$  and SAP, which is given as follows,

$$g_i = \sqrt{F_i^{-1}} h_i e^{j\phi}, \quad (2)$$

where  $h_i$  denotes the small scale fading coefficient of  $U_i$ , which follows the PDF of shadowed-Rician fading channel [41],  $e^{j\phi}$  represents the Doppler shift, which is caused by the movement of the LEO satellite. Assume that the diameter of a spot beam is about 100 km and the minimum elevation angle is  $30^\circ$  [42], we can calculate the average Doppler shift for all UEs is less than 1300 kHz and the maximum difference of Doppler shift between UEs is 15 kHz. To handle the Doppler shift on the LEO satellite channel, we can first pre-compensate the Doppler shift roughly by the predictable orbit and movement of satellite [43]. Then, a refine pre-compensation of the oscillator offsets can be performed by utilizing a phase-locked loop to track the reference tone [44]. In this way, the Doppler shift estimation can be controlled within a small error range. Further, to compensate for the residual frequency offset, specific modulation and demodulation techniques are required. For example, the cyclic prefix (CP) based method can be utilized to estimate the Doppler shift within  $\pm 1/2$  subcarrier for the OFDM system [45], [46]. In this way, the residual frequency offset can be estimated successfully in the most SNR regime. Moreover, as the total bandwidth of a Ka-band SAP is usually larger than 400 MHz [36], and the total frequency bandwidth is divided into multiple units with each channel bandwidth of  $B_c$ , we can set a guard band double than the Doppler shifts, denoted by  $B_g$ , to relieve the influence of Doppler shifts on the system performance [47], [48].  $F_i$  represents the free space path loss between SAP and  $U_i$ , which can be expressed as

$$F_i = 92.4 + 20 \log f + 20 \log d, \quad (3)$$

where  $f$  is the signal frequency and  $d$  is the distance between SAP and  $U_i$ .

### B. LSP-GFMA Scheme

Our LSP-GFMA scheme is a 2-step grant free random access scheme. The activated UEs randomly shift their selected rLPS to alleviate pilot collision and superimpose their sLPS on data to transmit. Then, the SAP performs ICEK and SIC decoding algorithm to recover the data of UEs. The details of LSP-GFMA scheme are shown as follows:

**Step 0:** The SAP configures uplink resources and broadcasts to the activated UEs, and also with its position and velocity for synchronization.

**Step 1:** In a certain time slot, the activated  $U_i$  randomly selects a rLPS  $\mathbf{p}_i$  from the LPS pilot set and randomly shifts  $\mathbf{p}_i$  by  $\tau_i$  bits. Then, the sLPS  $\mathbf{p}_{i,\tau_i} \in \mathbf{p}_{i,:} = \{\mathbf{p}_{i,\tau_1}, \mathbf{p}_{i,\tau_2}, \dots, \mathbf{p}_{i,\tau_K}\}$  is superimposed on the data  $\mathbf{d}_i$  at a specific power allocation ratio  $\alpha$ . Then the superimposed signal is sent to the SAP [49]. We assume that  $U_i$  is equipped with global navigation satellite system (GNSS) receiver to estimate the timing and carrier frequency for synchronization.

**Step 2:** The SAP knows the rLPS set and obtains all sLPS by shifting operation for each rLPS, and the SAP utilizes

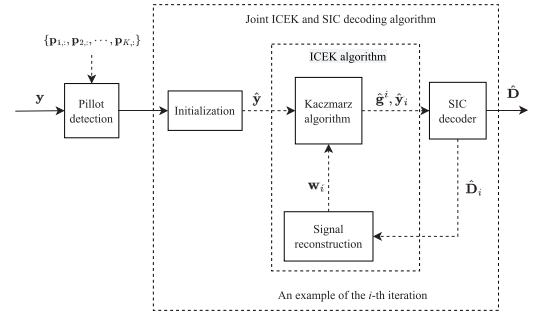


Fig. 2. In the procedure of joint ICEK and SIC decoding algorithm, after initialization, the decoded data by SIC decoder is reconstructed to re-estimate the CSI.

the sLPS  $\{\mathbf{p}_{1,:}, \mathbf{p}_{2,:}, \dots, \mathbf{p}_{K,:}\}$  to detect pilot collision and recover the data from  $\mathbf{y}$ . If a pilot collision occurs on  $\mathbf{p}_{i,\tau_i}$ , it would be discarded. Then, the SAP performs the joint ICEK and SIC decoding algorithm to recover the data of remaining UEs in  $\hat{\mathbf{y}}$  as shown in Fig. 2. In the  $i$ -th iteration, the ICEK algorithm outputs the estimated CSI  $\hat{\mathbf{g}}^i$  and  $\hat{\mathbf{y}}_i$ , and then delivers to the SIC decoder. If the algorithm converges, we get the decoded data  $\hat{\mathbf{D}} = \hat{\mathbf{D}}_i$ . Otherwise, the ICEK algorithm reconstructs the weighting matrix of superimposed signal  $\mathbf{W}_i$ , and updates CSI  $\hat{\mathbf{g}}^{i+1}$  with  $\mathbf{W}_i$  for the next iteration, the detailed process is introduced in Section III-A.

### III. DESIGN OF JOINT ICEK AND SIC DECODING ALGORITHM AND CONSTRUCTION OF LPS SET

In this section, we propose a joint ICEK and SIC decoding algorithm, and introduce the construction method of LPS set. Moreover, we prove that the average auto-correlation and cross-correlation of LPS are  $O(1/L)$ .

#### A. Joint ICEK and SIC Decoding Algorithm

In the SP structure, since the accurate CSI cannot be obtained due to the inherent data interference, and the channel estimation error leads to imperfect SIC in the decoding procedure, we propose a joint ICEK and SIC decoding algorithm.

The original Kaczmarz algorithm is known as algebraic reconstruction technique (ART) [50] in computed tomography, and it provides an iterative method for solving over-determined linear equation  $\mathbf{A}\mathbf{x} = \mathbf{b}$ , where  $\mathbf{A}$  is a  $n \times m$  matrix,  $\mathbf{x}$  is a  $m \times 1$  vector to be determined, and  $\mathbf{b}$  is a measurement vector. In every iterative process of Kaczmarz algorithm,  $\mathbf{A}_k$  is the  $k$ -th row of the matrix  $\mathbf{A}$ . The solution of the last inner iteration is  $\hat{\mathbf{x}}_{t,m-1}$  within the  $t$ -th outer iteration, which satisfies  $\langle \mathbf{A}_m^T, \hat{\mathbf{x}}_{t,m-1} \rangle = b_m$ . Given an initial solution  $\hat{\mathbf{x}}_{1,0}$ , the  $k$ -th inner iteration in the  $t$ -th outer iteration of Kaczmarz algorithm can be expressed as:

$$\begin{cases} \hat{\mathbf{x}}_{1,1} = \hat{\mathbf{x}}_{1,0} + \frac{b_1 - \langle \mathbf{A}_1^T, \hat{\mathbf{x}}_{1,0} \rangle}{\|\mathbf{A}_1\|^2} \mathbf{A}_1^T, \\ \dots \\ \hat{\mathbf{x}}_{t,k} = \hat{\mathbf{x}}_{t,k-1} + \frac{b_k - \langle \mathbf{A}_k^T, \hat{\mathbf{x}}_{t,k-1} \rangle}{\|\mathbf{A}_k\|^2} \mathbf{A}_k^T, \\ \text{for } t = 1, \dots, n, \text{ and } k = 1, \dots, m, \end{cases} \quad (4)$$

where  $\|\mathbf{A}_k\|$  is the second order norm of  $\mathbf{A}_k$ .

In general, if we utilize the MMSE estimation to obtain the CSI of activated UEs, Eq. (1) can be expressed as:

$$\mathbf{y} = (\mathbf{D} + \mathbf{P}_p)\mathbf{g} + \boldsymbol{\omega}, \quad (5)$$

where  $\mathbf{P}_p = \{\sqrt{(1-\alpha)P_t}\mathbf{p}_{1,\tau_1}, \dots, \sqrt{(1-\alpha)P_t}\mathbf{p}_{K,\tau_K}\}$  consists of sLPS,  $\mathbf{D} = \{\sqrt{\alpha P_t}\mathbf{d}_1, \sqrt{\alpha P_t}\mathbf{d}_2, \dots, \sqrt{\alpha P_t}\mathbf{d}_K\}$ , and  $\mathbf{g} = \{g_1, g_2, \dots, g_K\}^T$  is channel gain vector of  $K$  activated UEs. Let  $\mathbf{X} = \mathbf{D} + \mathbf{P}_p$ , and the estimated CSI of activated UEs can be expressed as:

$$\hat{\mathbf{g}} = \mathbf{W}\hat{\mathbf{y}} = (\mathbf{X}^H\mathbf{X} + \sigma^2\mathbf{I}_K)^{-1}\mathbf{X}^H\mathbf{y}, \quad (6)$$

where  $\mathbf{X}^H\mathbf{X}$  is a  $K \times K$  matrix,  $\mathbf{I}_K$  denotes a  $K \times K$  identity matrix,  $\mathbf{W} = (\mathbf{X}^H\mathbf{X} + \sigma^2\mathbf{I}_K)^{-1}$  represents the weighting matrix.  $\hat{\mathbf{y}} = \mathbf{P}_p^H\mathbf{y}$  can be utilized as a matched filter based on sLPS, which is expressed as:

$$\mathbf{P}_p^H\mathbf{y} = \mathbf{P}_p^H\mathbf{D}\mathbf{g} + \mathbf{P}_p^H\mathbf{P}_p\mathbf{g} + \mathbf{P}_p^H\boldsymbol{\omega}. \quad (7)$$

Note that  $\mathbf{P}_p^H\mathbf{P}_p$  is much larger than  $\mathbf{P}_p^H\mathbf{D}$  due to the low auto-correlation and cross-correlation of sLPS. Therefore, we can treat the data matrix  $\mathbf{D}$  as interference to the pilot matrix  $\mathbf{P}_p$  in our ICEK algorithm. To avoid the matrix inversion operation, we convert Eq. (6) to a linear equation  $\mathbf{W}^{-1}\hat{\mathbf{g}} = \hat{\mathbf{y}}$ , where  $\hat{\mathbf{y}}$  is regarded as a measurement vector, and the solution  $\hat{\mathbf{g}}$  can be calculated by the Kaczmarz algorithm with  $K$  inner iteration and  $J_1$  outer iteration:

$$\hat{\mathbf{g}}_{t,k} = \hat{\mathbf{g}}_{t,k-1} + \delta \frac{\hat{y}_k - \langle \mathbf{w}_k^T, \hat{\mathbf{g}}_{t,k-1} \rangle}{\|\mathbf{w}_k\|^2} \mathbf{w}_k^T, \quad (8)$$

where  $\hat{\mathbf{g}}_{t,k}$  is the solution in  $k$ -th inner iteration of  $t$ -th outer iteration,  $\mathbf{w}_k$  is the  $k$ -th row of  $\mathbf{W}^{-1}$ , which depends on the cross-correlation of selected sLPS by  $U_k$  and that of other UEs. The convergence factor  $\delta \in (1, 2)$  is usually set as  $(1 + \lambda_{\min}/\lambda_{\max})$ , where  $\lambda_{\min}$  and  $\lambda_{\max}$  represent the maximum and minimum eigenvalues of the matrix  $\mathbf{W}$ , respectively [51]. Consider  $\mathbf{W}$  is approximately diagonal matrix with the same elements and has  $\lambda_{\min} \approx \lambda_{\max}$  due to the low auto-correlation and cross-correlation of sLPS, we have  $\delta \approx 2$ . Furthermore, the initial solution is denoted as  $\hat{\mathbf{g}}_{1,0} = \mathbf{0}$ . Then, we should find the closest vector  $\hat{\mathbf{g}}_{t,k+1}$  to the true solution  $\mathbf{g}$  that satisfies

$$\arg \min_{\hat{\mathbf{g}}_{t,k-1}} |\langle \mathbf{w}_k^T, \hat{\mathbf{g}}_{t,k-1} \rangle - y_k|. \quad (9)$$

Therefore, we can find a large  $\|\mathbf{w}_k\|$  to obtain  $\langle \mathbf{w}_k^T, \hat{\mathbf{g}}_{t,k-1} \rangle = y_k$ . Consider that  $\|\mathbf{w}_k\|$  is determined by the cross-correlation of selected sLPS by  $U_k$  and that of other UEs, and  $\|\mathbf{w}_k\|$  increases with  $\sum_{i \neq k}^K |\tau_k - \tau_i|$ . Thus, we sort  $\|\mathbf{w}_k\|$  in a descending order as  $\{\|\mathbf{w}_{\psi_1}\| \geq \|\mathbf{w}_{\psi_2}\| \geq \dots \geq \|\mathbf{w}_{\psi_K}\|\}$  to accelerate the convergence to  $\langle \mathbf{w}_k^T, \hat{\mathbf{g}}_{t,k-1} \rangle = y_k$ . Then, Eq. (8) can be rewritten as:

$$\hat{\mathbf{g}}_{t,k} = \hat{\mathbf{g}}_{t,k-1} + \delta \frac{\hat{y}_{\psi_k} - \langle \mathbf{w}_{\psi_k}^T, \hat{\mathbf{g}}_{t,k-1} \rangle}{\|\mathbf{w}_{\psi_k}\|^2} \mathbf{w}_{\psi_k}^T. \quad (10)$$

Note that the estimated CSI from Eq. (10) needs further iteration process, because we only use  $\mathbf{P}_p$  to construct weighting matrix  $\mathbf{W}$  in ICEK algorithm yet. We can jointly utilize ICEK and SIC decoder, and iteratively update the CSI for improving the accuracy. In the first iteration of joint ICEK

---

**Algorithm 1** Joint ICEK and SIC Decoding Algorithm
 

---

**Input:**  $\mathbf{y}, \mathbf{P}_p, J_2, K, \hat{\mathbf{g}}_{1,0} = \mathbf{0}$   
**Output:**  $\hat{\mathbf{g}}$

- 1  $\hat{\mathbf{D}}_0 = \mathbf{0}, \hat{\mathbf{g}}^0 = \mathbf{0}, j = 1, t = 1, k = 1;$
- 2  $\mathbf{X}_0 = (1 - \alpha)\mathbf{P}_p + \alpha\hat{\mathbf{D}}_0, \hat{\mathbf{y}}_0 = \mathbf{X}_0^H\mathbf{y},$   
 $\mathbf{W}_0 = (\mathbf{X}_0^H\mathbf{X}_0 + \sigma_0^2\mathbf{I}_K);$
- 3  $\delta = (1 + \lambda_{\min}/\lambda_{\max});$
- 4 **for**  $j = 1, \dots, J_2$  **do**
- 5     **for**  $t = 1, \dots, J_1^{(j)}$  **do**
- 6         **for**  $k = 1, \dots, K$  **do**
- 7              $\hat{\mathbf{g}}_{t,k}^j = \hat{\mathbf{g}}_{t,k-1}^j + \delta \frac{y_{\psi_k} - \langle \mathbf{w}_{\psi_k}^T, \hat{\mathbf{g}}_{t,k-1}^j \rangle}{\|\mathbf{w}_{\psi_k}\|_2^2} \mathbf{w}_{\psi_k}^T;$
- 8             **end**
- 9             **if**  $\langle \mathbf{w}_{\psi_K}^T, \hat{\mathbf{g}}_{t,K-1}^j \rangle = y_{\psi_K}$  **then**
- 10                 **break;**
- 11             **end**
- 12     **end**
- 13      $\hat{\mathbf{D}}_{j+1} = \mathcal{F}(\hat{\mathbf{y}}_j, \hat{\mathbf{g}}^j);$  % perform SIC to decode data [9], [56];
- 14      $\mathbf{X}_{j+1} = \mathbf{P}_p + \hat{\mathbf{D}}_j;$
- 15      $\mathbf{W}_{j+1} = \mathbf{X}_{j+1}^H\mathbf{X}_{j+1} + \sigma_0^2\mathbf{I}_K;$
- 16 **end**

---

and SIC decoding algorithm, let  $\mathbf{X}_0 = \mathbf{P}_p + \hat{\mathbf{D}}_0$  and  $\hat{\mathbf{g}}^0 = \mathbf{0}$ , where  $\hat{\mathbf{D}}_0 = \mathbf{0}$ . Then, we construct the initial weighting matrix as  $\hat{\mathbf{W}}_0 = (\mathbf{X}_0^H\mathbf{X}_0 + \sigma^2\mathbf{I}_K)^{-1}$ , and utilize the ICEK algorithm to solve the linear equation  $\hat{\mathbf{W}}_0^{-1}\hat{\mathbf{g}}^0 = \hat{\mathbf{y}}$ . Note that  $J_1$  is the number of outer iterations when the ICEK algorithm achieves convergence, i.e.,  $\langle \mathbf{w}_{\psi_K}^T, \hat{\mathbf{g}}_{J_1, K-1}^j \rangle = y_{\psi_K}$ , and the number of outer iterations of ICEK in the  $j$ -th iteration of joint ICEK and SIC decoding algorithm is  $J_1^{(j)}$ .

Then, we can get  $\hat{\mathbf{g}}^1 = \hat{\mathbf{g}}_{J_1^{(1)}, K}^1$  after  $K$  inner iterations and  $J_1^{(1)}$  outer iterations, and deliver the estimated CSI  $\hat{\mathbf{g}}^1$  to the SIC decoder, denoted by  $\mathcal{F}$ . Note that the SIC performs in an iterative manner and stops when any UE fails to decode before all non-collision UEs are decoded [9], [56]. The first decoding result  $\hat{\mathbf{D}}_1 = \mathcal{F}(\mathbf{y}, \hat{\mathbf{g}}^1)$  is used to construct the superimposed signal matrix  $\mathbf{X}_1 = \mathbf{P}_p + \hat{\mathbf{D}}_1$  and  $\hat{\mathbf{W}}_1 = (\mathbf{X}_1^H\mathbf{X}_1 + \sigma^2\mathbf{I}_K)$ , where  $\mathcal{F}(\mathbf{y}, \hat{\mathbf{g}}^1)$  represents the data matrix decoded by SIC decoder with input  $\mathbf{y}$  and  $\hat{\mathbf{g}}^1$ . Then,  $\hat{\mathbf{W}}_1$  is the weighting matrix in the next iteration. Therefore, in the  $j$ -th iteration of joint ICEK and SIC decoding algorithm for  $j = 1, \dots, J_2$ , we can obtain the weighting matrix  $\hat{\mathbf{W}}_j = \mathbf{X}_j^H\mathbf{X}_j + \sigma_0^2\mathbf{I}_K$ , where  $\mathbf{X}_j = \mathbf{P}_p + \hat{\mathbf{D}}_j$ , and  $\hat{\mathbf{D}}_j = \mathcal{F}(\mathbf{y}, \hat{\mathbf{g}}^j)$ . In the ICEK algorithm, the CSI is updated iteratively to reduce the interference and then improve the estimation accuracy. In addition, the complexity of obtaining the weighting matrix in each iteration is  $\mathcal{O}(K^2)$ . Let  $J = \sum_{j=1}^{J_2} J_1^{(j)}$  denote the total number of iterations required to reach convergence. Thus, the ICEK algorithm has a combined complexity of  $\mathcal{O}(JK^2)$ . The detailed joint ICEK and SIC decoding algorithm is summarized in Algorithm 1, and the proof of convergence for line 9 in Algorithm 1 can be found in [52].

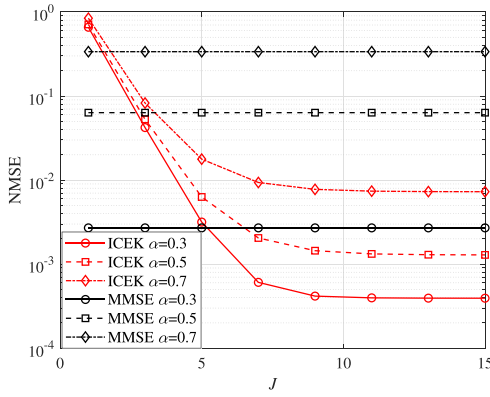


Fig. 3. NMSE performance of ICEK and MMSE algorithms with different  $\alpha$  versus  $J$  iterations.

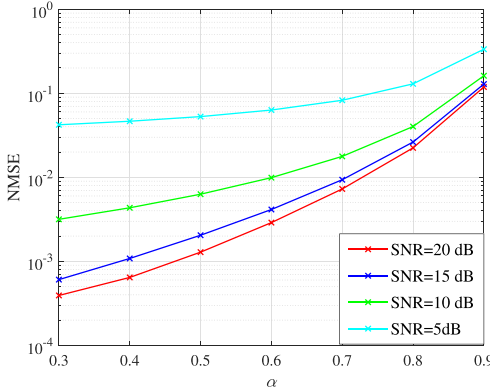


Fig. 4. NMSE of the ICEK algorithm versus  $\alpha$  with different SNR.

Besides, the power allocation coefficient  $\alpha$  has significant effect on the performance of joint ICEK and SIC decoder algorithm. To quantify the performance of joint ICEK and SIC decoding algorithm, we define the NMSE as follows,

$$NMSE = \frac{E(\Delta \mathbf{g}^H \Delta \mathbf{g})}{E(\mathbf{g}^H \mathbf{g})}, \quad (11)$$

where  $\Delta \mathbf{g} = |\hat{\mathbf{g}}_{J_2} - \mathbf{g}|$  denotes the channel estimation error.

Fig. 3 compares the NMSE of our ICEK and the conventional MMSE algorithms versus  $J$  iterations under different  $\alpha$ . We can observe that the ICEK algorithm achieves convergence after  $J = 10$  iterations for all the simulated  $\alpha$ , and it can achieve about 60 times gain over MMSE with respect to NMSE in our LSP-GFMA scheme. Moreover, the complexity of ICEK algorithm is  $O(JK^2)$ , while the complexity of MMSE algorithm is  $O(K^3)$ .

Furthermore, we can observe that the NMSE increases with the increase of  $\alpha$ . This is because the more power is allocated to data, the stronger data interference is led to the pilot, and we can improve the performance of ICEK algorithm by allocating more power to pilot. However, if  $\alpha$  is low, the SINR of UEs would be degraded and then the decoding failure probability would be deteriorated, which is analyzed in Section IV. Therefore, the selection of  $\alpha$  is a tradeoff between the accuracy of ICEK algorithm and SINR of UEs. We simulated the NMSE of our ICEK algorithm versus  $\alpha$  with

different SNR as shown in Fig. 4. The NMSE increases slowly as  $\alpha$  decreases when  $\alpha \leq 0.7$ .

### B. Construction of LPS Set

Both pilot detection and joint ICEK and SIC decoding algorithm are based on LPS set, which plays an important role in our LSP-GFMA scheme.

We can construct the LPS set by utilizing the Hadamard matrix  $\mathbf{H} = [h_{i,j}]$  of order  $M$  with elements  $\{-1, 1\}$ , and a length  $N$  perfect sequence  $\mathbf{a} = \{a_1, a_2, \dots, a_{N-1}\}$  with elements  $\{-1, 0, 1\}$ . Note that for a zero-correlation-zone periodic sequences set with zero-correlation zone  $[0, \tau_z]$  [53], each pair of sequences in this set has zero cross-correlation if they left shift no more than  $\tau_z$  bits symbols. We expand the zero-correlation zone to the low-correlation zone to construct our LPS set as follows.

**Step I:** Let  $M = M' \times N'$  and rearrange each row  $\mathbf{h}_{i,:} = \{h_{i,0}, h_{i,1}, \dots, h_{i,(M-1)}\}$  of  $\mathbf{H}$  to a  $M' \times N'$  matrix, and we have

$$\mathbf{H}_i = \begin{bmatrix} h_{0,0}^{(i)} & h_{0,1}^{(i)} & \cdots & h_{0,(N'-1)}^{(i)} \\ h_{1,0}^{(i)} & h_{1,1}^{(i)} & \cdots & h_{1,(N'-1)}^{(i)} \\ \vdots & \vdots & \vdots & \vdots \\ h_{(M'-1),0}^{(i)} & h_{(M'-1),1}^{(i)} & \cdots & h_{(M'-1),(N'-1)}^{(i)} \end{bmatrix},$$

where  $h_{m,n}^{(i)} = h_{i,m \times N' + n}$ ,  $0 \leq m \leq M' - 1$  and  $0 \leq n \leq N' - 1$ .  $\mathbf{H}_i$  can be represented as  $\{\mathbf{h}_{0,:}^{(i)}, \mathbf{h}_{1,:}^{(i)}, \dots, \mathbf{h}_{m,:}^{(i)}, \dots, \mathbf{h}_{M'-1,:}^{(i)}\}$ , where  $\mathbf{h}_{m,:}^{(i)}$  represents a row vector in the  $m$ -th row of  $\mathbf{H}_i$ .

**Step II:** Let  $\mathbf{s} = \{s_0, s_1, \dots, s_i, \dots, s_{N'-1}\}$  be a shift sequence of length  $N'$ , where  $s_i$  is an integer and  $s_i \in [0, N]$ . We can obtain the interleaving matrix  $\mathbf{I}$  with the length  $N$  perfect sequence  $\mathbf{a}$  and  $\mathbf{s}$ :

$$\mathbf{I} = \begin{bmatrix} a_{0+s_0} & a_{0+s_1} & \cdots & a_{0+s_{N'-1}} \\ a_{1+s_0} & a_{1+s_1} & \cdots & a_{1+s_{N'-1}} \\ \vdots & \vdots & \vdots & \vdots \\ a_{N-1+s_0} & a_{N-1+s_1} & \cdots & a_{N-1+s_{N'-1}} \end{bmatrix}, \quad (12)$$

where the subscript is calculated by modulo  $N$ . For simplicity, the interleaving matrix can be expressed as

$$\mathbf{I} = \{\mathbf{L}^{s_0}(\mathbf{a}), \mathbf{L}^{s_1}(\mathbf{a}), \dots, \mathbf{L}^{s_{N'-1}}(\mathbf{a})\}, \quad (13)$$

where  $\mathbf{L}^s(\cdot)$  denotes the operator of  $s$  symbols left shift. Then, for  $0 \leq i \leq M - 1$ , we construct a matrix  $\mathbf{P}^{(i)}$  with  $N \times M'$  rows and  $N'$  columns,

$$\mathbf{P}^{(i)} = \begin{bmatrix} \mathbf{I} \circ \mathbf{h}_{0,:}^{(i)} \\ \mathbf{I} \circ \mathbf{h}_{1,:}^{(i)} \\ \vdots \\ \mathbf{I} \circ \mathbf{h}_{M'-1,:}^{(i)} \end{bmatrix}, \quad (14)$$

where “ $\circ$ ” is the column-wise product, which is defined as follows:

$$\begin{aligned} \mathbf{I} \circ \mathbf{h}_{m,:}^{(i)} &= \{\mathbf{L}^{s_0}(\mathbf{a}) \cdot h_{m,0}^{(i)}, \mathbf{L}^{s_1}(\mathbf{a}) \cdot h_{m,1}^{(i)}, \dots, \\ &\quad \times \mathbf{L}^{s_{N'-1}}(\mathbf{a}) \cdot h_{m,(N'-1)}^{(i)}\}. \end{aligned} \quad (15)$$

**Step III:** By concatenating the successive rows of the matrix  $\mathbf{P}^{(i)}$ , it can be rearranged to a rLPS sequence  $\mathbf{p}_i = \{p_0^{(i)}, p_1^{(i)}, \dots, p_{N \times M - 1}^{(i)}\}$  with length  $N \times M$ . Then, we can get the rLPS set  $\mathbf{P} = \{\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_{M-1}\}$ .

Moreover, we can set  $s_i = \lfloor \frac{N}{N'} \rfloor \times i \pmod{N}$  to generate optimal LPS set [53]. In this condition, the bound of zero-correlation zone is  $\tau_z = N' \times \lfloor \frac{N}{N'} \rfloor - 1$ , and the maximum symbols that LPS can left shift is  $\tau_m = N \times N' - 1$  to guarantee the low correlation. Thus, the length of LPS is  $L = N \times M$ . The detail derivation is given in the Section III-C.

A toy example to better understand the LPS is given as follows:

For a perfect sequence  $\mathbf{a} = \{1, 1, 0, 1, 0, 0, -1\}$  with length  $N = 7$  and a Hadamard matrix  $\mathbf{H}$  with order  $M = 4$ , can be expressed as

$$\mathbf{H} = \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \end{bmatrix}. \quad (16)$$

Let  $M = M' \times N'$ , and  $M' = N' = 2$ . The shift sequence with  $N' = 2$  is calculated as  $\mathbf{s} = \{0, 3\}$  via  $s_i = \lfloor \frac{N}{N'} \rfloor \times i \pmod{N}$  in Step II. Then, the toy example to generate the LPS set according to the above steps can be summarized as follows.

**Step I,** we rearrange each row of  $\mathbf{H}$  to a  $M' \times N'$  matrix, and we have:

$$\begin{aligned} \mathbf{H}_0 &= \begin{bmatrix} +1 & +1 \\ +1 & +1 \end{bmatrix}, & \mathbf{H}_1 &= \begin{bmatrix} +1 & -1 \\ +1 & -1 \end{bmatrix}, \\ \mathbf{H}_2 &= \begin{bmatrix} +1 & +1 \\ -1 & -1 \end{bmatrix}, & \mathbf{H}_3 &= \begin{bmatrix} +1 & -1 \\ -1 & +1 \end{bmatrix}. \end{aligned} \quad (17)$$

**Step II,** we construct an interleaving matrix with  $\mathbf{a} = \{1, 1, 0, 1, 0, 0, -1\}$  and  $\mathbf{s} = \{0, 3\}$  according to Eq. (12) as follows:

$$\mathbf{I} = \begin{bmatrix} +1 & +1 & 0 & +1 & 0 & 0 & -1 \\ +1 & 0 & 0 & -1 & +1 & +1 & 0 \end{bmatrix}^T, \quad (18)$$

where the first column of  $\mathbf{I}$  is obtained by left shifting  $\mathbf{a} = \{1, 1, 0, 1, 0, 0, -1\}$  with  $s_0 = 0$ , and the second column of  $\mathbf{I}$  is obtained by left shifting  $\mathbf{a} = \{1, 1, 0, 1, 0, 0, -1\}$  with  $s_1 = 3$ .

Let  $\mathbf{I} = \{\mathbf{I}_0, \mathbf{I}_1\}$ , where  $\mathbf{I}_0 = \{+1, +1, 0, +1, 0, 0, -1\}^T$  and  $\mathbf{I}_1 = \{+1, 0, 0, -1, +1, +1, 0\}^T$ . Moreover, let  $\mathbf{H}_i = \{\mathbf{h}_{0,:}^{(i)}, \mathbf{h}_{1,:}^{(i)}\}$ . Then, we can construct a matrix  $\mathbf{P}^{(i)}$  with  $N \times M' = 7 \times 2$  rows and  $N' = 2$  columns as follows:

$$\mathbf{P}^{(i)} = \begin{bmatrix} \mathbf{I} \circ \mathbf{h}_{0,:}^{(i)} \\ \mathbf{I} \circ \mathbf{h}_{1,:}^{(i)} \end{bmatrix}. \quad (19)$$

Therefore, we can obtain  $\mathbf{P}^{(0)}, \mathbf{P}^{(1)}, \mathbf{P}^{(2)}$  and  $\mathbf{P}^{(3)}$  as follows:

$$\begin{aligned} \mathbf{P}^{(0)} &= \begin{bmatrix} +1 & +1 \\ +1 & 0 \\ 0 & 0 \\ +1 & -1 \\ 0 & +1 \\ 0 & +1 \\ -1 & 0 \\ +1 & +1 \\ +1 & 0 \\ 0 & 0 \\ +1 & -1 \\ 0 & +1 \\ 0 & +1 \\ -1 & 0 \end{bmatrix}, & \mathbf{P}^{(1)} &= \begin{bmatrix} +1 & -1 \\ +1 & 0 \\ 0 & 0 \\ +1 & +1 \\ 0 & -1 \\ 0 & -1 \\ -1 & 0 \\ +1 & -1 \\ +1 & 0 \\ 0 & 0 \\ +1 & +1 \\ 0 & -1 \\ 0 & -1 \\ -1 & 0 \end{bmatrix}, \\ \mathbf{P}^{(2)} &= \begin{bmatrix} +1 & +1 \\ +1 & 0 \\ 0 & 0 \\ +1 & -1 \\ 0 & +1 \\ 0 & +1 \\ -1 & 0 \\ -1 & -1 \\ -1 & 0 \\ 0 & 0 \\ -1 & +1 \\ 0 & -1 \\ 0 & -1 \\ +1 & 0 \end{bmatrix}, & \mathbf{P}^{(3)} &= \begin{bmatrix} +1 & -1 \\ +1 & 0 \\ 0 & 0 \\ +1 & +1 \\ 0 & -1 \\ 0 & -1 \\ -1 & 0 \\ -1 & +1 \\ -1 & 0 \\ 0 & 0 \\ -1 & -1 \\ 0 & +1 \\ 0 & +1 \\ +1 & 0 \end{bmatrix}. \end{aligned} \quad (20)$$

**Step III,** we successively concatenate the rows of  $\mathbf{P}^{(i)}$  ( $i = 0, 1, 2, 3$ ) and rearrange them to be rLPS sequences  $\mathbf{p}_i$  as follows:

$$\begin{aligned} \mathbf{p}_0 &= \{+1, +1, +1, 0, 0, 0, +1, -1, 0, +1, 0, +1, -1, 0, \\ &\quad +1, +1, +1, 0, 0, 0, +1, -1, 0, +1, 0, +1, -1, 0\}, \\ \mathbf{p}_1 &= \{+1, -1, +1, 0, 0, 0, +1, +1, 0, -1, 0, -1, -1, 0, \\ &\quad +1, -1, +1, 0, 0, 0, +1, +1, 0, -1, 0, -1, -1, 0\}, \\ \mathbf{p}_2 &= \{+1, +1, +1, 0, 0, 0, +1, -1, 0, +1, 0, +1, -1, 0, \\ &\quad -1, -1, -1, 0, 0, 0, -1, +1, 0, -1, 0, -1, +1, 0\}, \\ \mathbf{p}_3 &= \{+1, -1, +1, 0, 0, 0, +1, +1, 0, -1, 0, -1, -1, 0, \\ &\quad -1, +1, -1, 0, 0, 0, -1, -1, 0, +1, 0, +1, +1, 0\}. \end{aligned} \quad (21)$$

Finally, we can get the rLPS set  $\mathbf{P} = \{\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$  with  $M = 4$ ,  $L = M \times N = 28$ ,  $\tau_z = N' \times \lfloor \frac{N}{N'} \rfloor - 1 = 5$  and  $\tau_m = N \times N' - 1 = 13$ .

### C. Theoretical Analysis of LPS

In this subsection, we prove that if the number of left shift symbols is  $\tau \in [0, \tau_z]$  on rLPS, each pair of sLPS in a LPS set are orthogonal to each other. Moreover, if the number of left shift symbols is  $\tau \in (\tau_z, \tau_m]$ , any pair of sLPS in this LPS set are non-orthogonal with a low average cross-correlation  $O(1/L)$ . Let  $R_{\mathbf{p}_{i,\tau_i}, \mathbf{p}_{j,\tau_j}}$  denote the cross-correlation function between two sLPS  $\mathbf{p}_{i,\tau_i}$  and  $\mathbf{p}_{j,\tau_j}$  with

their left shift difference is  $\Delta\tau = |\tau_i - \tau_j|$  symbols, and we have

$$R_{\mathbf{p}_{i,\tau_i}, \mathbf{p}_{j,\tau_j}}(\Delta\tau) = \frac{1}{L} \sum_{n=0}^{L-1} p_{\tau_i+n}^{(i)} \cdot p_{\tau_j+n}^{(j)}. \quad (22)$$

Thus, we have the following *Theorem 1* and *Theorem 2* to prove that both of the auto-correlation and cross-correlation are 0 and  $O(1/L)$  in  $\Delta\tau \in [0, \tau_z]$  and  $\Delta\tau \in (\tau_z, \tau_m]$ , respectively.

**Theorem 1:** If  $0 \leq \Delta\tau \leq \tau_z$ , we always have  $R_{\mathbf{p}_{i,\tau_i}, \mathbf{p}_{j,\tau_j}}(\tau) = 0$  for arbitrary two sLPS  $\mathbf{p}_{i,\tau_i}$  and  $\mathbf{p}_{j,\tau_j}$ , where the condition  $i = j$  and  $\tau_i = \tau_j$  are not simultaneously true.

*Proof:* Let  $\Delta\tau = \tau_q \times N' + \tau_r$ , where  $0 \leq \tau_q \leq \lfloor \frac{N'}{N'} \rfloor - 1$ ,  $0 \leq \tau_r \leq N' - 1$ . According to [54], the cross-correlation between  $\mathbf{p}_{i,\tau_i}$  and  $\mathbf{p}_{j,\tau_j}$  can be calculated as

$$\begin{aligned} R_{\mathbf{p}_{i,\tau_i}, \mathbf{p}_{j,\tau_j}}(\Delta\tau) &= \frac{1}{L} \sum_{k=0}^{M'-1} \left( \sum_{t=0}^{N'-\tau_r-1} \left( h_{k,t}^{(i)} h_{k,(t+\tau_r)}^{(j)} \right) R_{\mathbf{a}}(s_{t+\tau_r} - s_t + \tau_q) \right. \\ &\quad \left. + \sum_{t=N'-\tau_r}^{N'-1} \left( h_{k,t}^{(i)} h_{k,(t+\tau_r)}^{(j)} \right) R_{\mathbf{a}}(s_{t+\tau_r} - s_t + \tau_q + 1) \right), \end{aligned} \quad (23)$$

where  $R_{\mathbf{a}}(\Delta\tau)$  denotes the auto-correlation of perfect sequences  $\mathbf{a}$ , and  $R_{\mathbf{a}}(\Delta\tau) = 0$  if  $\Delta\tau \bmod N \neq 0$ . Then, we can distinguish the following two conditions:

- 1) When  $0 < \Delta\tau \leq \tau_z$ , two cases can be further distinguished: If  $0 \leq t \leq N' - \tau_r - 1$ , we have  $s_{t+\tau_r} - s_t + \tau_q \geq 0$ ; Else if  $N' - \tau_r \leq t \leq N' - 1$ , we have  $s_{t+\tau_r} \geq s_t$ , then we have  $s_{t+\tau_r} - s_t = \tau_r \times \lfloor \frac{N'}{N'} \rfloor$ , thus we have  $s_{t+\tau_r} - s_t + \tau_q > 0$  and  $s_{t+\tau_r} - s_t + \tau_q \leq N' \times \lfloor \frac{N'}{N'} \rfloor - 1 < N$ ; Therefore, we have  $R_{\mathbf{p}_{i,\tau_i}, \mathbf{p}_{j,\tau_j}}(\Delta\tau) = 0$ .
- 2) When  $\Delta\tau = 0$  and  $i \neq j$ , we have  $R_{\mathbf{p}_{i,\tau_i}, \mathbf{p}_{j,\tau_j}}(0) = R_{\mathbf{a}}(0) \sum_{k=0}^{M'-1} \sum_{t=0}^{N'-1} (h_{k,t}^{(i)} h_{k,t}^{(j)}) = 0$ .

Combining the above two conditions, we finish the proof of *Theorem 1*.  $\square$

Therefore, the LPS has perfect auto-correlation and cross-correlation if the number of left shift symbols is in  $[0, \tau_z]$ . Note that the average cross-correlation of LPS affects the decoding failure probability in massive access, and we prove that the average auto-correlation and cross-correlation of LPS are  $O(1/L)$  in the following *Theorem 2*.

**Theorem 2:** If  $\tau_z < \Delta\tau \leq \tau_m$ , the average auto-correlation and cross-correlation of arbitrary two sLPS  $\mathbf{p}_{i,\tau_i}$  and  $\mathbf{p}_{j,\tau_j}$  are  $O(1/L)$ .

*Proof:* Let  $\Delta\tau = \tau_q \times N' + \tau_r, \lfloor \frac{N'}{N'} \rfloor \leq \tau_q < N$ ,  $0 \leq \tau_r \leq N' - 1$ .

For  $0 \leq t \leq N' - \tau_r - 1$ , we have  $(t + \tau_r) \bmod N' - t \geq 0$ , i.e.,  $s_{t+\tau_r} - s_t \geq 0$ , then  $s_{t+\tau_r} - s_t + \tau_q > 0$ .

Since  $s_{t+\tau_r} - s_t + \tau_q \leq (N' - 1) \lfloor \frac{N'}{N'} \rfloor + (N - 1) < 2N$ , we only have  $R_{\mathbf{a}}(s_{t+\tau_r} - s_t + \tau_q) \neq 0$  if  $s_{t+\tau_r} - s_t + \tau_q = N$ .

Then, we can directly find that  $R_{\mathbf{a}}(s_{t+\tau_r} - s_t + \tau_q + 1) = 0$ .

Similarly, for  $N' - \tau_r \leq t \leq N' - 1$ , we have  $s_{t+\tau_r} - s_t \leq 0$  and  $-N < s_{t+\tau_r} - s_t + \tau_q < N$ , and we only have  $R_{\mathbf{a}}(s_{t+\tau_r} - s_t + \tau_q) \neq 0$  if  $s_{t+\tau_r} - s_t + \tau_q = 0$ .

Note that the length of LPS is  $L = N \times M$ , and  $M = M' \times N'$  is the order of Hadamard matrix to construct the LPS set. If  $\tau_z < \Delta\tau \leq \tau_m$ , the average auto-correlation and cross-correlation of arbitrary two sLPS  $\mathbf{p}_{i,\tau_i}$  and  $\mathbf{p}_{j,\tau_j}$  are

$$\begin{aligned} \bar{R}_{\mathbf{p}_{i,\tau_i}, \mathbf{p}_{j,\tau_j}} &= \frac{1}{L} \times \frac{1}{N \times N' - \tau_z} \times \sum_{\tau=\tau_z}^{\tau_m} R_{\mathbf{p}_{i,\tau_i}, \mathbf{p}_{j,\tau_j}}(\Delta\tau) \\ &= \frac{1}{L} \times \frac{1}{N \times N' - \tau_z} \times \sum_{k=0}^{M'-1} \left( \sum_{t=0}^{N'-\tau_r-1} \left( h_{k,t}^{(i)} h_{k,(t+\tau_r)}^{(j)} \right) \right. \\ &\quad \left. \times R_{\mathbf{a}}(N) + \sum_{t=N'-\tau_r}^{N'-1} \left( h_{k,t}^{(i)} h_{k,(t+\tau_r)}^{(j)} \right) R_{\mathbf{a}}(0) \right) \\ &\leq \frac{1}{L} \cdot \frac{N' \times M' \times N}{N \times N' - \tau_z} \leq \frac{M}{N' - 1} \cdot \frac{1}{L}. \end{aligned} \quad (24)$$

Therefore, the average auto-correlation and cross-correlation of arbitrary two sLPS are  $O(1/L)$  when the left symbol shift is in  $(\tau_z, \tau_m]$ , which is lower than that of mZCS  $O(1/\sqrt{L})$ .  $\square$

#### IV. PERFORMANCE ANALYSIS OF LSP-GFMA SCHEME WITH LPS UNDER IMPERFECT SIC

In this section, we assume that the pilot collision can be detected with signal power detections [55] or utilize the CRC [56]. We derive the close-form expressions of the AFP and achievable throughput of the uplink LSP-GFMA under shadowed-Rician fading channel and imperfect SIC. AFP is defined as the probability that pilot collision occurs or the data of UEs fails to be decoded. The achievable throughput is defined as the number of successfully decoded data bits in each data frame.

Note that for  $K$  activated UEs randomly select rLPS pilots, the pilot collision probability  $\beta$  can be calculated as:

$$\beta = 1 - \left( 1 - \frac{1}{M \times (\tau_m + 1)} \right)^{K-1}, \quad (25)$$

where  $M \times (\tau_m + 1) = M \times N' \times N$  is the total number of sLPS. After the pilot detection, the UEs in pilot collision are discarded. Assume the number of UEs without pilot collision is  $K_s$ , and the probability distribution of  $K_s$  can be expressed as:

$$\Pr(K_s) = \binom{K}{K_s} (1 - \beta)^{K_s} (\beta)^{K-K_s}. \quad (26)$$

Then, if  $U_i$  selects  $\mathbf{p}_i$  and shifts  $\tau_i$ , the SAP can recover data  $\hat{\mathbf{d}}_i$  with sLPS  $\mathbf{p}_{i,\tau_i}$  from  $K_s$  UEs without pilot collision to obtain  $\mathbf{d}_i$  as follows,

$$\begin{aligned} \hat{\mathbf{d}}_i &= \frac{(\mathbf{p}_{i,\tau_i})^T}{\|\mathbf{p}_{i,\tau_i}\|^2} \frac{1}{\sqrt{\alpha P_t}} \mathbf{y} \\ &= \mathbf{g}_i \frac{(\mathbf{p}_{i,\tau_i})^T \mathbf{p}_{i,\tau_i}}{\|\mathbf{p}_{i,\tau_i}\|^2} \left( \mathbf{d}_i + \sqrt{\frac{1-\alpha}{\alpha}} \mathbf{p}_{i,\tau_i} \right) + \frac{(\mathbf{p}_{i,\tau_i}^T \mathbf{p}_{j,\tau_j})}{\|\mathbf{p}_{i,\tau_i}\|^2} \\ &\quad \times \sum_{j=1}^{K_n} \mathbf{g}_j \left( \mathbf{d}_j + \sqrt{\frac{1-\alpha}{\alpha}} \mathbf{p}_{j,\tau_j} \right) + \frac{(\mathbf{p}_{i,\tau_i})^T}{\|\mathbf{p}_{i,\tau_i}\|^2} \frac{1}{\sqrt{\alpha P_t}} \omega, \end{aligned} \quad (27)$$



where the UEs that select orthogonal sLPS to  $\mathbf{p}_{i,\tau_i}$  are eliminated, and  $K_n$  is the number of UEs that select low-correlation sLPS to  $\mathbf{p}_{i,\tau_i}$ , which determines the interference of low-correlation UEs to  $U_i$ . The probability distribution of  $K_n$  can be expressed as:

$$\begin{aligned} \Pr(K_n) &= \frac{\binom{K_s-1}{K_n} (M(\tau_z + 1) - 1)^{K_s-1-K_n} (M\tau_m - M\tau_z)^{K_n}}{(M(\tau_m + 1) - 1)^{K_s-1}}, \end{aligned} \quad (28)$$

where  $K_n = 0, 1, \dots, K_s - 1$ . Note that we can utilize typical threshold-based algorithms to detect UEs, and define the false alarm probability as the probability of detecting a sequence when it has not been transmitted. This happens when the correlator output is larger than a predetermined threshold value. The threshold is set as the estimation of the interference of low-correlation UEs to  $U_i$  in Eq. (27) and noise power  $\sigma^2$ , i.e.,  $\frac{P_t(K_n-1)}{LK_n} \cdot \sum_{i=1}^{K_n} \mathbf{g}_i + \sigma^2$ . Thus, the false alarm probability is lower than  $10^{-4}$  under all simulation parameters, which can be ignored in our analysis. Then, the decoding failure probability of  $U_i$  can be calculated as:

$$\xi = 1 - \sum_{K_s=0}^K \Pr(K_s) \sum_{K_n=0}^{K_s-1} \Pr(K_n) \Pr(\gamma \geq \gamma^{th} | K_n), \quad (29)$$

where  $\gamma$  represents SINR of  $U_i$ . Note that we ignore the specific modulation and channel coding schemes, and utilize the decoding threshold  $\gamma^{th}$  to judge whether a decoding process can be successful, i.e., if  $\gamma \geq \gamma^{th}$ ,  $\mathbf{d}_i$  can be recovered successfully.  $\gamma^{th}$  is dependent on the modulation and channel coding schemes that decide the data transmission rate [18], [32]. Considering imperfect SIC, the SINR of  $U_i$  is calculated as:

$$\gamma = \frac{\alpha\theta|g_i|^2}{(1-\alpha)\theta|g_i|^2 + I + \epsilon}, \quad (30)$$

where  $\theta = \frac{P_t}{\sigma^2}$  represents SNR, and  $I$  represents interference from other low-correlation UEs,

$$I = \mu \sum_{t=0}^{K_h} \frac{\mathbf{p}_{i,\tau_i}^T \mathbf{p}_{t,\tau_t}}{\|\mathbf{p}_{i,\tau_i}\|^2} \theta |g_i|^2 + \sum_{v=K_h+1}^{K_n-1} \frac{\mathbf{p}_{i,\tau_i}^T \mathbf{p}_{v,\tau_v}}{\|\mathbf{p}_{i,\tau_i}\|^2} \theta |g_v|^2, \quad (31)$$

where  $K_h$  denotes the number of decoded UEs.  $\mu$  is interference transfer factor of imperfect SIC [35].

Further,  $\epsilon = \sum_{t=0}^{K_h} |\Delta g_i|^2 + 1$ , where  $\sum_{t=0}^{K_h} |\Delta g_i|^2$  is the interference caused by channel estimation error. The definition expression of  $\Delta g_i$  is heuristic as shown in Eq. (10), and we set  $\Delta g_i$  as a constant from simulation results.

Denote  $\mathcal{X} = \theta|g_i|^2$ , the PDF of  $\mathcal{X}$  over the shadowed-Rician fading channel can be expressed as:

$$\begin{aligned} f_{\mathcal{X}}(x) &= \left( \frac{2b_0m_0}{2b_0m_0 + \Omega} \right)^{m_0} \frac{1}{2b_0\theta} \exp\left(-\frac{x}{2b_0\theta}\right) \\ &\times {}_1F_1\left(m_0, 1, \frac{\Omega x}{2b_0\theta(2b_0m_0 + \Omega)}\right), x \geq 0, \end{aligned} \quad (32)$$

where  $2b_0$  and  $m_0$  represent the average power of the multipath component and Nakagami-m parameter, respectively.  $\Omega$  denotes line-of-sight (LoS) component, and  ${}_1F_1\left(m_0, 1, \frac{\Omega x}{2b_0\theta(2b_0m_0 + \Omega)}\right)$  denotes the confluent hypergeometric function, which can be expressed as:

$$\begin{aligned} {}_1F_1\left(m_0, 1, \frac{\Omega x}{2b_0\theta(2b_0m_0 + \Omega)}\right) &= \\ \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\Gamma(m_0 + n)}{\Gamma(n + 1)} \left( \frac{\Omega x}{2b_0\theta(2b_0m_0 + \Omega)} \right)^n. \end{aligned} \quad (33)$$

We select the parameters in [41] as  $b_0 = 0.063$ ,  $m_0 = 1$ ,  $\Omega = 0.000897$ , and the PDF can be simplified as:

$$\begin{aligned} f_{\mathcal{X}}(x) &= \left( \frac{2b_0m_0}{2b_0m_0 + \Omega} \right)^{m_0} \frac{1}{2b_0\theta} \exp\left(-\frac{x}{2b_0\theta}\right) \\ &\times \exp\left(\frac{\Omega x}{2b_0\theta(2b_0m_0 + \Omega)}\right)^n \\ &= \rho e^{-\eta x}, \quad 0 \leq x < +\infty, \end{aligned} \quad (34)$$

where  $\rho = \frac{1}{2b_0\theta} \left( \frac{2b_0m_0}{2b_0m_0 + \Omega} \right)^{m_0}$ ,  $\eta = \frac{m_0}{(2b_0m_0 + \Omega)\theta}$ . Then,  $I$  can be expressed as  $\sum_{j=0}^{K_n-1} \rho_j \mathcal{X}_j$ , where

$$\rho_j = \begin{cases} \mu \frac{\mathbf{p}_{i,\tau_i}^T \mathbf{p}_{j,\tau_j}}{\|\mathbf{p}_{i,\tau_i}\|^2}, & 0 \leq j \leq K_h, \\ \frac{\mathbf{p}_{i,\tau_i}^T \mathbf{p}_{j,\tau_j}}{\|\mathbf{p}_{i,\tau_i}\|^2}, & K_h < j \leq K_n - 1. \end{cases} \quad (35)$$

The PDF of  $I$  can be derived as:

$$f_I(y) = \sum_{j=0}^{K_n} \frac{\rho}{\rho_j} e^{-\frac{\eta}{\rho_j} y} \prod_{l \neq j} \frac{\frac{\rho}{\rho_l}}{\left(\frac{\rho}{\rho_l} - \frac{\rho}{\rho_j}\right)}, \quad 0 \leq y < +\infty. \quad (36)$$

The detail derivation can be found in Appendix A.

Since  $\mathcal{X}$  and  $I$  are independent, the joint PDF of  $\mathcal{X}$  and  $I$  is

$$\begin{aligned} f(x, y) &= f_{\mathcal{X}}(x) f_I(y). \\ &= \sum_{j=0}^{K_n-1} \frac{\rho^2}{\rho_j} e^{-\eta\left(\frac{y}{\rho_j} + x\right)} \prod_{l \neq j} \frac{\frac{\rho}{\rho_l}}{\left(\frac{\rho}{\rho_l} - \frac{\rho}{\rho_j}\right)}, \end{aligned} \quad (37)$$

where  $0 \leq x < +\infty$  and  $0 \leq y < +\infty$ . Then, when  $K_n$  UEs select low-correlation sLPS to  $\mathbf{p}_{i,\tau_i}$  under  $\gamma^{th}$ , the decoding success probability  $\Pr(\gamma \geq \gamma^{th} | K_n)$  can be calculated as follows

$$\begin{aligned} \Pr(\gamma \geq \gamma^{th} | K_n) &= \frac{1}{K_n} \sum_{K_h=0}^{K_n-1} \Pr\left(\frac{\alpha \mathcal{X}}{(1-\alpha)\mathcal{X} + I + \epsilon} \geq \gamma^{th}\right) \\ &= \frac{1}{K_n} \sum_{K_h=0}^{K_n-1} \Pr\left(\mathcal{X} \geq \frac{I + \epsilon}{\left(\frac{\alpha}{\gamma^{th}} - 1 + \alpha\right)}\right) \\ &= \frac{1}{K_n} \sum_{K_h=0}^{K_n-1} \Pr(\mathcal{X} \geq \phi I + \phi \epsilon) \\ &= \frac{1}{K_n} \sum_{K_h=0}^{K_n-1} \sum_{j=0}^{K_n-1} \frac{\rho^2}{\rho_j} \prod_{l \neq j} \frac{\frac{\rho}{\rho_l}}{\left(\frac{\rho}{\rho_l} - \frac{\rho}{\rho_j}\right)} \end{aligned}$$

$$\begin{aligned} & \times \int_0^{+\infty} dy \int_{\phi y + \phi \epsilon}^{+\infty} e^{-\eta(\frac{y}{\rho_j} + x)} dx, \\ & = \frac{1}{K_n} \sum_{K_h=0}^{K_n-1} \sum_{j=0}^{K_n-1} \frac{\rho^2 e^{\phi \epsilon}}{\eta(\phi \rho_j + 1)} \prod_{l \neq j} \frac{\frac{\rho}{\rho_l}}{\left(\frac{\rho}{\rho_l} - \frac{\rho}{\rho_j}\right)}, \end{aligned} \quad (38)$$

where  $\phi = \frac{1}{\left(\frac{\alpha}{\gamma^{th}} - 1 + \alpha\right)}$ .

Therefore, we can substitute Eq. (26), Eq. (28) and Eq. (38) into Eq. (29), and the decoding failure probability can be expressed as,

$$\begin{aligned} \xi & = 1 - \sum_{K_s=0}^{K-1} \Pr(K_s) \sum_{K_n=0}^{K_s-1} \sum_{K_h=0}^{K_n-1} \sum_{j=0}^{K_n-1} \frac{1}{K_n} \\ & \times \binom{K-1}{K_n} \left( \frac{M(\tau_m + 1) + \tau_z - \tau_m - 1}{M(\tau_m + 1) - 1} \right)^{K-1-K_n} \\ & \times \left( \frac{\tau_m - \tau_z}{M(\tau_m + 1) - 1} \right)^{K_n} \frac{\rho^2 e^{\phi \epsilon}}{\eta(\phi \rho_j + 1)} \prod_{l \neq j} \frac{\frac{\rho}{\rho_l}}{\left(\frac{\rho}{\rho_l} - \frac{\rho}{\rho_j}\right)}. \end{aligned} \quad (39)$$

Finally, we can calculate AFP from Eq. (25) and Eq. (39) as follows,

$$\zeta = 1 - (1 - \beta) \times (1 - \xi). \quad (40)$$

The expression of achievable throughput can be derived by utilizing AFP as follows,

$$\zeta = B \times K \times (1 - \zeta). \quad (41)$$

## V. SIMULATION RESULTS AND DISCUSSION

In this section, we simulate the pilot collision probability, decoding failure probability and AFP of LSP-GFMA scheme with LPS, ZCS, and mZCS. Then, we compare AFP and achievable throughput of LSP-GFMA scheme with the existing GFMA schemes, including the modified approximate message passing (M-AMP) [8] and the E-SSA schemes [57] with mZCS and ZCS, respectively. Both the M-AMP and E-SSA schemes adopt the RP structure. Note that the length of ZCS  $L_{ZCS}$  is a prime number, and the mZCS is generated from  $L_{ZCS} - 1$  different roots. As the largest number of available orthogonal pilot sequences is equal to the pilot length, we try to construct the LPS with similar length to the ZCS for fair comparison of the pilot collision probability and decoding failure probability. Thus, in the simulation setting, we set the length of ZCS as 173 and 337, and the length of LPS as 168, 336. The simulation parameters are listed in Table I [58], [59].

Fig. 5(a) shows the pilot collision probability of three different pilots with respect to  $K$  activated UEs with different length of pilots. We can observe that the LPS achieves about  $M$  times lower pilot collision probability than that of ZCS, and the mZCS has the lowest pilot collision probability because it has the most number of non-orthogonal sequences from  $L_{ZCS} - 1$  different roots. However, the decoding failure probability of mZCS is significantly deteriorated with the increasing of  $K$  due to its high cross-correlation than the LPS.

Fig. 5(b) shows the decoding failure probability of LSP-GFMA scheme with LPS, mZCS and ZCS pilots with respect

TABLE I  
SIMULATION PARAMETERS

Parameters	Value
Transmit Power $P_t$	[100 mW, 10W]
Transmit Distance $d$	1200 km
Transmit Frequency $f$	30 GHz
Guard Bandwidth $B_g$	180 kHz
Total Bandwidth	400 MHz
Subcarrier Bandwidth $B_c$	200 kHz
Maximum Transmit Antenna Gain	43.2 dBi
Maximum Satellite Antenna Gain	51dBi
Interference Transfer Factor $\mu$	0.0001
Activation Probability $p_a$	0.01
Thermal Noise Density	-174 dBm/Hz
LPS Pilot Length $L_{LPS}$	168, 336
Order of Hadamard Matrix $M$	8, 16
Length of Perfect Sequence $N$	21
Length of Shift Sequence $N'$	8, 16
ZCS Pilot Length $L_{ZCS}$	173, 337
Decoding Threshold $\gamma^{th}$	1.2
Pilot Allocation Coefficient $\alpha$	0.7

to the decoding threshold  $\gamma^{th}$ . Under the same SNR, ZCS has the lowest decoding failure probability, because the uncollission UEs with ZCS are zero cross-correlation to each other. Due to the low cross-correlation  $O(1/L)$  of LPS, it has about 23 times lower decoding failure probability than mZCS. To find the most appropriate  $\gamma^{th}$ , we define  $G_\xi = \frac{\xi_{LPS}^2}{\xi_{mZCS} \xi_{ZCS}}$  as the gain of LPS over mZCS and ZCS for decoding failure probability  $\xi$ , and achieve the largest  $G_\xi=1.625$  for  $\gamma^{th} = 1.2$  when  $K = 50$ . Therefore, we set  $\gamma^{th} = 1.2$  in the following simulations.

Note that the time difference between the UEs at sub-satellite point and the edge of spot beam will lead to the pilot unalignment at the satellite receiver, and the decoding threshold  $\gamma^{th}$  also becomes larger. With our simulation setup, there may one sampling point unalignment and we find that  $\gamma^{th} = 1.25$  can achieve the largest  $G_\xi$ . Thus, Fig. 5(c) shows the decoding failure probability with and without time difference versus SNR. It can be observed that the decoding failure probability of LPS with time difference is slightly higher than that of LPS without time difference when SNR  $< 10$ dB. Moreover, the time difference increases the decoding failure probability by less than 1% when SNR  $\geq 10$ dB. Hence, it is reasonable to assume that the impact of the time difference on the proposed scheme can be ignored when an appropriate SNR is guaranteed.

Fig. 6(a) shows the decoding failure probability with respect to  $K$  activated UEs under different  $\gamma^{th}$  and  $L$ , which also validates the accuracy of Eq. (39). Note that the uncollission UEs with ZCS are zero cross-correlation to each other, which lead to the decoding failure probability of ZCS is not affected by  $K$ . However, both of the UEs with LPS or mZCS may suffer the interference from other uncollission UEs, and the decoding failure probability increases with  $K$ . We can observe that LPS has a lower decoding failure probability than mZCS, because the UEs with mZCS suffer higher interference due

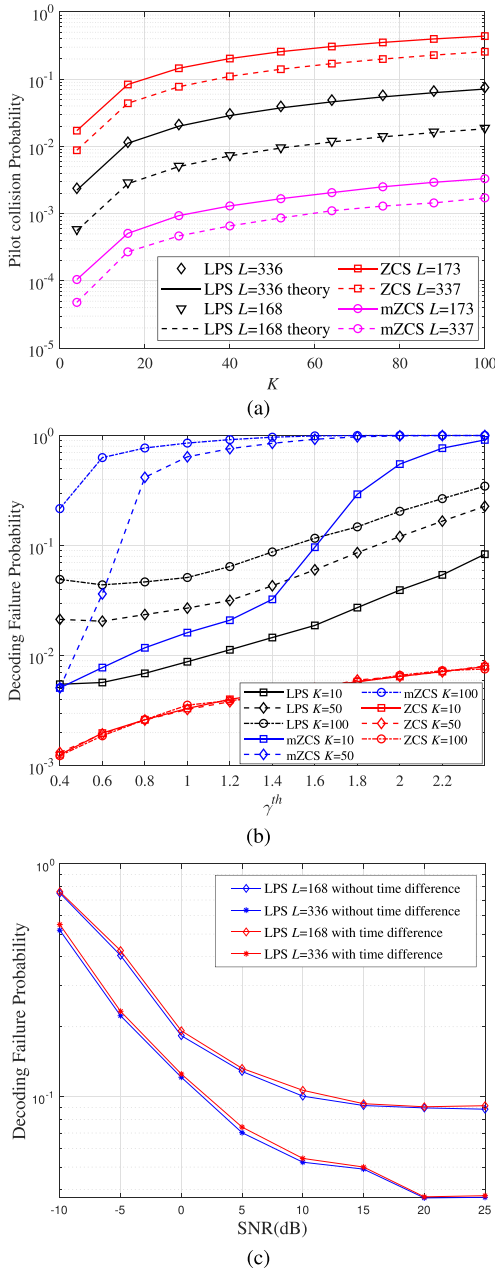


Fig. 5. Pilot collision probability and decoding failure probability: (a) Pilot collision probability versus  $K$  for LPS with  $L_{LPS} = 168, 336$ , ZCS with  $L_{ZCS} = 173, 337$  and mZCS with  $L_{ZCS} = 173, 337$ , (b) Decoding failure probability versus  $\gamma^{th}$  with different  $K$ , where SNR = 10 dB and  $\alpha = 0.7$ , (c) Decoding failure probability of LPS with and without time difference versus SNR, where  $K = 50$  and  $\alpha = 0.7$ .

to  $O(1/\sqrt{L})$  cross-correlation of mZCS, where that of LPS is  $O(1/L)$ .

Fig. 6(b) shows the decoding failure probability with respect to SNR under  $K = 50$  and different lengths of pilots. Similar to Fig. 6(a), the UEs with ZCS are zero cross-correlation to each other, which lead to the lowest decoding failure probability and decreases as the SNR increases. On the other hand, when  $SNR \geq 10$  dB, the decoding failure probability of UEs with LPS and mZCS are not sensitive to SNR, but mainly determined by the interference of superimposed pilot and number of other UEs. Moreover, the decoding failure probability of LPS with  $L_{LPS} = 336$  is about

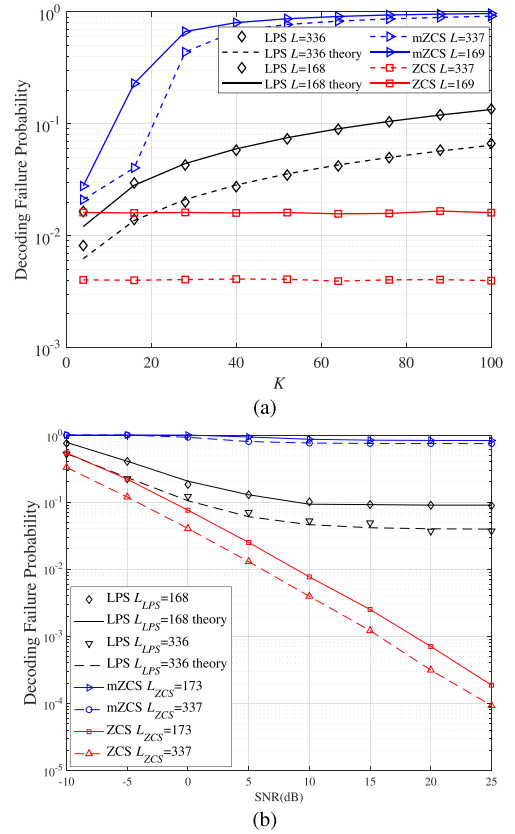


Fig. 6. Decoding failure probability: (a) Decoding failure probability versus  $K$  with different  $L$ , where SNR is 10 dB and  $\alpha = 0.7$ , (b) Decoding failure probability versus SNR with  $K = 50$ , where  $\gamma^{th} = 1.2$  and  $\alpha = 0.7$ .

10 times lower than mZCS with  $K = 50$  when they are stable as  $SNR \geq 10$  dB.

Fig. 7(a) shows the AFP of LSP-GFMA scheme versus  $\alpha$  under different SNR. It can be observed that the AFP of LSP-GFMA scheme in all SNR regimes can achieve a minimum value when  $\alpha=0.7$ . This is because smaller  $\alpha$  leads to higher channel estimation accuracy. However, when  $\alpha$  is small, the SINR of UEs would be degraded and then the AFP is deteriorated.

The comparison of AFP of three pilots versus  $K$  under different  $L$  and  $\gamma^{th}$  is shown in Fig. 7(b), where the simulation results validate the accuracy of Eq. (41). We can observe that the AFP of LPS is lower than ZCS in all  $K$  and  $L$  regimes, because the AFP of ZCS is seriously deteriorated by the pilot collision due to its limited size of pilot set as shown in Fig. 4. On contrast, the AFP of UEs with LPS and mZCS are mainly affected by the decoding failure probability, which are both limited by the interference from other UEs. Therefore, the AFP of UEs with LPS can be lower than that of UEs with mZCS, which validates that our LSP-GFMA scheme with the innovative designed LPS can achieve better tradeoff between the pilot collision probability and decoding failure probability than both ZCS and mZCS pilots.

Fig. 8(a) compares the AFP of three pilots with different SNR under different length of pilots with  $K = 50$ . The AFP of LPS is the lowest among three pilots under all SNRs when  $K = 50$ , which indicates that the LSP-GFMA scheme with LPS has a great advantage in practical massive access. When

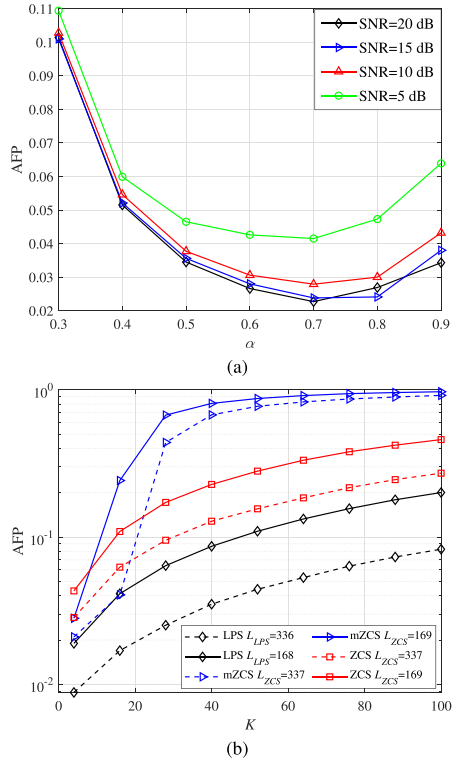


Fig. 7. AFP of LSP-GFMA scheme: (a) AFP versus  $\alpha$  with different SNR, where  $\gamma^{th} = 1.2$ ,  $L = 336$  and  $K = 50$ , (b) AFP versus  $K$  with different pilot length, where SNR is 10 dB,  $\alpha = 0.7$  and  $\gamma^{th} = 1.2$ .

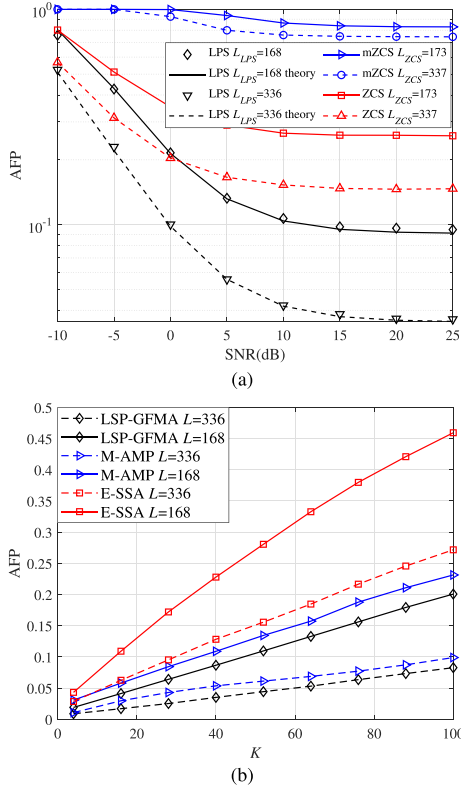


Fig. 8. AFP of LSP-GFMA and other schemes: (a) AFP versus SNR with  $K = 50$ , where  $\gamma^{th} = 1.2$  and  $\alpha = 0.7$ , (b) AFP of LSP-GFMA, E-SSA and M-AMP versus  $K$  with different  $L$  where SNR=10 dB.

$\text{SNR} \geq 10$  dB, the AFP of three pilots are not sensitive to SNR, where the AFP of ZCS is mainly determined by the pilot collision, and the AFP of LPS and mZCS are mainly affected

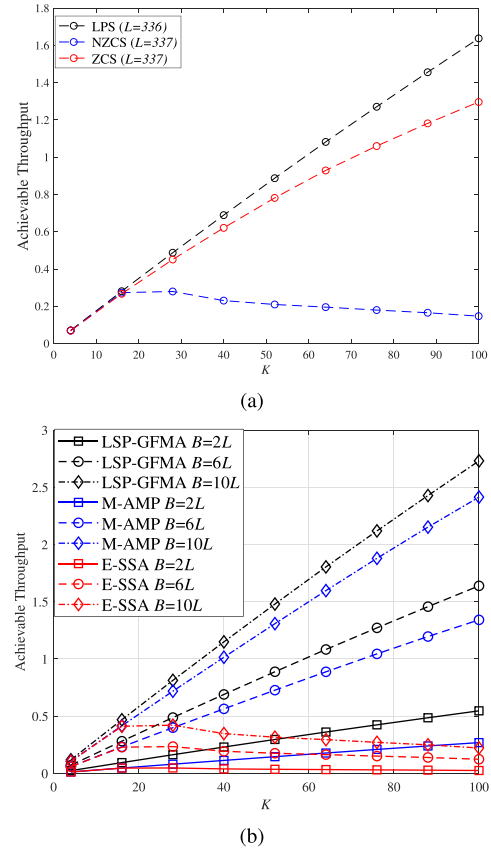


Fig. 9. Achievable throughput comparison: (a) Achievable throughput of LPS, mZCS and ZCS versus  $K$ , where  $B = 6L$ , (b) Achievable throughput of LSP-GFMA, E-SSA and M-AMP versus  $K$  with different  $L$  where SNR=10 dB.

by the decoding failure probability due to the interference of other UEs. Moreover, the AFP of LPS with  $L_{LPS} = 336$  is about 60 times lower than that of mZCS and 10 times lower than that of ZCS when  $\text{SNR} \geq 10$  dB.

Fig. 8(b) shows the AFP of LSP-GFMA, E-SSA and M-AMP schemes versus  $K$ . We can observe that our LSP-GFMA has the lowest AFP among three schemes. In the E-SSA scheme, the activated UEs adopt the same ZCS pilot with different delay, and the SAP can distinguish them by different delays. The number of different delays is equal to the length of ZCS pilot, which leads to the pilot collision probability of E-SSA is higher than our LSP-GFMA, and thereby the AFP of E-SSA is about 2.5 times higher than our LSP-GFMA scheme. In addition, the M-AMP scheme utilizes the sparsity to reduce the correlation between pilots based on compressed sensing algorithm. The shorter pilot leads to higher correlation [8], and the correlation of adopted pilot set in the M-AMP scheme is higher than that of the LSP-GFMA scheme when pilot length is 336 or 168. Consequently, the M-AMP scheme has a slightly higher AFP than our LSP-GFMA scheme. Besides, the computational complexity of M-AMP scheme is  $O(K/p_a \cdot L)$ , where our LSP-GFMA scheme is  $O(JK^2)$ .

Fig. 9(a) compares the achievable throughput of three pilots with different number of active UEs. Recalling that the AFP of ZCS is mainly deteriorated by the pilot collision due to its limited size of pilot set, while the AFP of LPS and mZCS

are mainly affected by the decoding failure probability, which are both limited by the interference from other UEs. Then, the achievable throughput of LPS is higher than that of ZCS and mZCS when the pilot sequence length is almost the same. Finally, we compare the achievable throughput of our LSP-GFMA scheme with E-SSA and M-AMP schemes versus  $K$  with different length of data  $B$  as shown in Fig. 9(b). Similarly, the proposed LSP-GFMA scheme has the largest achievable throughput among three schemes, because the LSP-GFMA scheme adopts superimposed pilot structure, which does not take up additional time-frequency resources.

## VI. CONCLUSION

In this paper, we proposed an LSP-GFMA scheme with innovative designed LPS for uplink mMTC in S-IoT. We constructed the LPS set with  $M$  orthogonal rLPS by utilizing a perfect sequence and Hadamard matrix, and the activated UEs could randomly select and left shift the rLPS to obtain sLPS and significantly alleviate pilot collision. The pilot collision probability of our LSP-GFMA scheme with LPS is  $M$  times lower than that of conventional ZCS. We proved that sLPS has lower average auto-correlation and cross-correlation of  $O(1/L)$  than mZCS of  $O(1/\sqrt{L})$ , and derived the accurate theoretical expression of AFP and achievable throughput of the LSP-GFMA scheme. Moreover, we proposed a joint ICEK and SIC decoding algorithm for obtaining accurate CSI. Simulation results validated the accuracy of our derivations, and demonstrated that the decoding failure probability of LPS is about 60 times lower than that of mZCS when  $K \geq 50$ . Furthermore, the LSP-GFMA scheme with LPS design could achieve lower AFP than ZCS and mZCS with the increasing number of activated UEs  $K$ , which indicated that LPS has a great advantage in practical massive access. We also demonstrated that our LSP-GFMA scheme can achieve lower AFP and higher achievable throughput than the existing M-AMP and E-SSA schemes.

## APPENDIX A

### DERIVATION OF PDF OF INTERFERENCE FROM LOW-CORRELATION UES UNDER IMPERFECT SIC

Based on Eq. (31), we have  $\rho = \eta$  under the selected parameters. The interference  $I$  from low-correlation UEs can be expressed as:

$$I = \sum_{j=0}^{K_n-1} \rho_j \mathcal{X}_j, \quad (42)$$

where  $\rho_j = \mu \frac{p_{i,\tau_i}^T p_{j,\tau_j}^*}{\|p_{i,\tau_i}\|^2}$  when  $0 \leq j < K_h$ , and  $\rho_j = \frac{p_{i,\tau_i}^T p_{j,\tau_j}^*}{\|p_{i,\tau_i}\|^2}$  when  $K_h < j \leq K_n - 1$ .

Let  $\mathcal{Z}_j = \rho_j \mathcal{X}_j$  and  $I = \sum_{j=0}^{K_n-1} \mathcal{Z}_j$ . We can derive the PDF of  $\mathcal{Z}_j$  as:

$$f_{\mathcal{Z}_j}(z) = \frac{\rho}{\rho_j} e^{\frac{\rho}{\rho_j} z}, \quad 0 \leq z < +\infty. \quad (43)$$

The Laplace transform of  $\mathcal{Z}_j$  is as follows:

$$T(f_{\mathcal{Z}_j}(z)) = \frac{\frac{\rho}{\rho_j}}{s + \frac{\rho}{\rho_j}}. \quad (44)$$

Since  $\{\mathcal{Z}_0, \mathcal{Z}_1, \dots, \mathcal{Z}_{K_n-1}\}$  are independent, then we have

$$F_I(s) = T(f_I(z)) = \prod_{j=0}^{K_n-1} \frac{\rho}{\rho_j} \prod_{j=0}^{K_n-1} \left(s + \frac{\rho}{\rho_j}\right)^{-1}. \quad (45)$$

The function  $F_I(s)$  is a proper rational function, which can be represented as a sum of partial fractions in the following manner:

$$F_I(s) = \sum_{j=0}^{K_n-1} \frac{c_j}{\left(s + \frac{\rho}{\rho_j}\right)}, \quad (46)$$

where the constants  $c_j$  are given as follows:

$$\begin{aligned} c_j &= \lim_{s \rightarrow -\frac{\rho}{\rho_j}} F_I(s) \left(s + \frac{\rho}{\rho_j}\right) \\ &= \prod_{j=1}^{K_n-1} \frac{\rho}{\rho_j} \lim_{s \rightarrow -\frac{\rho}{\rho_j}} \prod_{\substack{l=0 \\ l \neq j}}^{K_n-1} \left(s + \frac{\rho}{\rho_l}\right) \\ &= \frac{\rho}{\rho_j} \prod_{\substack{l=0 \\ l \neq j}}^{K_n-1} \frac{\frac{\rho}{\rho_l}}{\left(\frac{\rho}{\rho_l} - \frac{\rho}{\rho_j}\right)}. \end{aligned} \quad (47)$$

Finally, we can derive the  $f_I(y)$  in Eq. (36) by utilizing the inverse Laplace transform:

$$\begin{aligned} f_I(y) &= T^{-1}(F_I(s)) = \sum_{j=0}^{K_n-1} c_j e^{-\mu_j y} \\ &= \sum_{j=0}^{K_n-1} \frac{\rho}{\rho_j} e^{-\frac{\rho}{\rho_j} y} \prod_{\substack{l=0 \\ l \neq j}}^{K_n-1} \frac{\frac{\rho}{\rho_l}}{\left(\frac{\rho}{\rho_l} - \frac{\rho}{\rho_j}\right)}, \quad 0 \leq y < +\infty. \end{aligned} \quad (48)$$

## REFERENCES

- [1] B. Di, L. Song, Y. Li, and H. V. Poor, "Ultra-dense LEO: Integration of satellite access networks into 5G and beyond," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 62–69, Apr. 2019.
- [2] J. Jiao, S. Wu, R. Lu, and Q. Zhang, "Massive access in space-based Internet of Things: Challenges, opportunities, and future directions," *IEEE Wireless Commun.*, vol. 28, no. 5, pp. 118–125, Oct. 2021.
- [3] M. Centenaro, C. E. Costa, F. Granelli, C. Sacchi, and L. Vangelista, "A survey on technologies, standards and open challenges in satellite IoT," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1693–1720, 3rd Quart., 2021.
- [4] T. Yang, J. Jiao, S. Wu, R. Lu, and Q. Zhang, "Grant free age-optimal random access protocol for satellite-based Internet of Things," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 3947–3961, Jun. 2022.
- [5] X. Zhou et al., "Active terminal identification, channel estimation, and signal detection for grant-free NOMA-OTFS in LEO satellite Internet-of-Things," *IEEE Trans. Wireless Commun.*, vol. 22, no. 4, pp. 2847–2866, Apr. 2023.
- [6] E. Casini, R. De Gaudenzi, and O. D. R. Herrero, "Contention resolution diversity slotted Aloha (CRDSA): An enhanced random access scheme for satellite access packet networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1408–1419, Apr. 2007.
- [7] O. Del Rio Herrero and R. De Gaudenzi, "High efficiency satellite multiple access scheme for machine-to-machine communications," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 4, pp. 2961–2989, Oct. 2012.
- [8] K. Senel and E. G. Larsson, "Grant-free massive MTC-enabled massive MIMO: A compressive sensing approach," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6164–6175, Dec. 2018.
- [9] M. Ke, Z. Gao, Y. Wu, X. Gao, and R. Schober, "Compressive sensing-based adaptive active user detection and channel estimation: Massive access meets massive MIMO," *IEEE Trans. Signal Process.*, vol. 68, pp. 764–779, 2020.

- [10] J. Ding, D. Qu, P. Liu, and J. Choi, "Machine learning enabled preamble collision resolution in distributed massive MIMO," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2317–2330, Apr. 2021.
- [11] J. Jiao, L. Xu, S. Wu, R. Lu, and Q. Zhang, "MSPA: Multislot pilot allocation random access protocol for mMTC-enabled IoT system," *IEEE Internet Things J.*, vol. 8, no. 24, pp. 17403–17416, Dec. 2021.
- [12] K. Lee, J. Kim, J. Jung, and I. Lee, "Zadoff–Chu sequence based signature identification for OFDM," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4932–4942, Oct. 2013.
- [13] J. Choi, J. Ding, N. P. Le, and Z. Ding, "Grant-free random access in machine-type communication: Approaches and challenges," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 151–158, Feb. 2022.
- [14] Z. Liu, J. Jiao, S. Wu, R. Lu, and Q. Zhang, "Non-orthogonal superimposed pilot grant-free random access scheme in satellite-based IoT," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2022, pp. 1407–1412.
- [15] S. Bisen, P. Shaik, and V. Bhatia, "On performance of energy harvested cooperative NOMA under imperfect CSI and imperfect SIC," *IEEE Trans. Veh. Technol.*, vol. 70, no. 9, pp. 8993–9005, Sep. 2021.
- [16] D. Verenzuela, E. Björnson, and L. Sanguinetti, "Spectral and energy efficiency of superimposed pilots in uplink massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7099–7115, Nov. 2018.
- [17] K. Upadhyaya, S. A. Vorobyov, and M. Vehkaperä, "Superimposed pilots are superior for mitigating pilot contamination in massive MIMO," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2917–2932, Jun. 2017.
- [18] *Service Requirements for the 5G System*, document TS 22.261, V18.6.0, 3GPP, 2022.
- [19] J. Ding, D. Qu, and J. Choi, "Analysis of non-orthogonal sequences for grant-free RA with massive MIMO," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 150–160, Jan. 2020.
- [20] G. Schreiber and M. Tavares, "5G new radio physical random access preamble design," in *Proc. IEEE 5G World Forum (5GWF)*, Jul. 2018, pp. 215–220.
- [21] R.-A. Pitaval, B. M. Popovic, P. Wang, and F. Berggren, "Overcoming 5G PRACH capacity shortfall: Supersets of Zadoff–Chu sequences with low-correlation zone," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5673–5688, Sep. 2020.
- [22] Y. Du et al., "Efficient multi-user detection for uplink grant-free NOMA: Prior-information aided adaptive compressive sensing perspective," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2812–2828, Dec. 2017.
- [23] A. T. Abebe and C. G. Kang, "Multi-sequence spreading random access (MSRA) for compressive sensing-based grant-free communication," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7531–7543, Nov. 2021.
- [24] B. Wang, L. Dai, Y. Zhang, T. Mir, and J. Li, "Dynamic compressive sensing-based multi-user detection for uplink grant-free NOMA," *IEEE Commun. Lett.*, vol. 20, no. 11, pp. 2320–2323, Nov. 2016.
- [25] S. Kim, J. Kim, and D. Hong, "A new non-orthogonal transceiver for asynchronous grant-free transmission systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1889–1902, Mar. 2021.
- [26] P. Bohlin and M. Tapio, "Performance evaluation of MIMO communication systems based on superimposed pilots," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2004, pp. 1–5.
- [27] H. Yin, D. Gesbert, and L. Cottatellucci, "Dealing with interference in distributed large-scale MIMO systems: A statistical approach," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 942–953, Oct. 2014.
- [28] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, Oct. 2014.
- [29] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.
- [30] M. Teeti, J. Sun, D. Gesbert, and Y. Liu, "The impact of physical channel on performance of subspace-based channel estimation in massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 4743–4756, Sep. 2015.
- [31] Y. Liu, H. Wang, W. Zhang, Q. Xu, and L. Shen, "Decoding method based on complex ICA for a multicell massive MIMO uplink system," *IEEE Signal Process. Lett.*, vol. 23, no. 5, pp. 648–652, May 2016.
- [32] F. Li, H. Wang, M. Ying, W. Zhang, and J. Lu, "Channel estimations based on superimposed pilots for massive MIMO uplink systems," in *Proc. 8th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2016, pp. 1–5.
- [33] J. Jiao, J. Zhou, S. Wu, and Q. Zhang, "Superimposed pilot code-domain NOMA scheme for satellite-based Internet of Things," *IEEE Syst. J.*, vol. 15, no. 2, pp. 2732–2743, Jun. 2021.
- [34] A. Wang, L. Lei, E. Lagunas, A. I. Pérez-Neira, S. Chatzinotas, and B. Ottersten, "Joint optimization of beam-hopping design and NOMA-assisted transmission for flexible satellite systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 10, pp. 8846–8858, Oct. 2022.
- [35] J. Zhao, X. Yue, S. Kang, and W. Tang, "Joint effects of imperfect CSI and SIC on NOMA based satellite-terrestrial systems," *IEEE Access*, vol. 9, pp. 12545–12554, 2021.
- [36] I. Leyva-Mayorga et al., "LEO small-satellite constellations for 5G and beyond-5G communications," *IEEE Access*, vol. 8, pp. 184955–184964, 2020.
- [37] *System Architecture for the 5G System*, document TS 23.501, V17.2.0, 3GPP, 2022.
- [38] *Solutions for NR to Support Non-Terrestrial Networks (NTN)*, document TS 38.821, V16.0.0, 3GPP, 2019.
- [39] Y. Polyanskiy, "A perspective on massive random-access," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 2523–2527.
- [40] K. Nikitopoulos, F. Mehran, and H. Jafarkhani, "Space-time super-modulation: Concept, design rules, and its application to joint medium access and rateless transmission," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 8275–8288, Dec. 2017.
- [41] A. Abdi, W. C. Lau, M. Alouini, and M. Kaveh, "A new simple model for land mobile satellite channels: First- and second-order statistics," *IEEE Trans. Wireless Commun.*, vol. 2, no. 3, pp. 519–528, May 2003.
- [42] E. Kim, I. P. Roberts, P. A. Iannucci, and J. G. Andrews, "Downlink analysis of LEO multi-beam satellite communication in shadowed Rician channels," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 1–6.
- [43] *Study on Narrow-Band Internet of Things (NB-IoT)/Enhanced Machine Type Communication (eMTC) Support for Non-Terrestrial Networks (NTN)*, document TR 36.763, V17.0.0, 3GPP, Jun. 2021.
- [44] K.-U. Storek, R. T. Schwarz, and A. Knopp, "Multi-satellite multi-user MIMO precoding: Testbed and field trial," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–7.
- [45] L. You, K.-X. Li, J. Wang, X. Gao, X.-G. Xia, and B. Ottersten, "Massive MIMO transmission for LEO satellite communications," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1851–1865, Aug. 2020.
- [46] J. Lin, Z. Hou, Y. Zhou, L. Tian, and J. Shi, "Map estimation based on Doppler characterization in broadband and mobile LEO satellite communications," in *Proc. IEEE 83rd Veh. Technol. Conf. (VTC Spring)*, May 2016, pp. 1–5.
- [47] *Solutions for NR to Support Non-Terrestrial Networks (NTN)*, document TR 38.821, V16.1.0, 3GPP, Jun. 2021.
- [48] D. Goto, H. Shibayama, F. Yamashita, and T. Yamazato, "LEO-MIMO satellite systems for high capacity transmission," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.
- [49] X. Lin et al., "Doppler shift estimation in 5G new radio non-terrestrial networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 1–6.
- [50] M. N. Boroujerdi, S. Haghghatshoar, and G. Caire, "Low-complexity statistically robust precoder/detector computation for massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6516–6530, Oct. 2018.
- [51] F. Natterer, *The Mathematics of Computerized Tomography*. New York, NY, USA: Wiley, 1986, pp. 107–108.
- [52] L. Dai and T. B. Schön, "On the exponential convergence of the Kaczmarz algorithm," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1571–1574, Oct. 2015.
- [53] P. Ke and Z. Zhou, "A generic construction of Z-periodic complementary sequence sets with flexible flock size and zero correlation zone length," *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1462–1466, Sep. 2015.
- [54] G. Golomb and G. Gong, *Signal Designs With Good Correlations: For Wireless Communications, Cryptography and Radar Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [55] L. Bai, J. Liu, Q. Yu, J. Choi, and W. Zhang, "A collision resolution protocol for random access in massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 686–699, Mar. 2021.
- [56] K. Zhang, J. Jiao, S. Wu, and Q. Zhang, "Short analog fountain code with quasi-gray constellation mapping modulation towards uRLC," *IEEE Trans. Signal Process.*, vol. 70, pp. 4077–4092, 2022.
- [57] A. Reichman, "Enhanced spread spectrum Aloha (E-SSA), an emerging satellite return link messaging scheme," in *Proc. IEEE 28th Conv. Electr. Engineers Isr. (IEEEI)*, Dec. 2014, pp. 1–4.

- [58] J. Sedin et al., “5G massive machine type communication performance in non-terrestrial networks with LEO satellites,” in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 01–06.
- [59] B. Di, H. Zhang, L. Song, Y. Li, and G. Y. Li, “Ultra-dense LEO: Integrating terrestrial-satellite networks into 5G and beyond for data offloading,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 47–62, Jan. 2019.



**Liang Xu** received the B.S. degree in communication engineering from Nanchang University, Nanchang, China, in 2018, and the M.S. degree in information and communication engineering from the Harbin Institute of Technology (Shenzhen), Shenzhen, China, in 2021, where he is currently pursuing the Ph.D. degree in information and communication engineering with the Guangdong Provincial Key Laboratory of Aerospace Communication and Networking Technology. His research interests include space communication protocols, massive access, age of information, and semantic communications.



**Jian Jiao** (Senior Member, IEEE) received the M.S. and Ph.D. degrees in communication engineering from the Harbin Institute of Technology (HIT), Harbin, China, in 2007 and 2011, respectively.

From 2011 to 2015, he was a Post-Doctoral Research Fellow with the Communication Engineering Research Centre, HIT Shenzhen (HITSz) Graduate School, Shenzhen, China. From 2016 to 2017, he was a China Scholarship Council Visiting Scholar with the School of Electrical and Information Engineering, The University of Sydney, Sydney, Australia. Since 2017, he has been with HITSz, where he has been a Professor with the Guangdong Provincial Key Laboratory of Aerospace Communication and Networking Technology since 2022. He is currently an Associate Professor with the Peng Cheng Laboratory, Shenzhen. His current research interests include semantic communications, satellite communications and networking, and coding techniques.

Prof. Jiao is also serving as an Editor for *Science China Information Sciences*.



**Ye Wang** (Member, IEEE) received the M.S. and Ph.D. degrees in information and communication engineering from the Harbin Institute of Technology (HIT), Harbin, China, in 2009 and 2013, respectively. From 2013 to 2014, he was a Post-Doctoral Research Fellow with the University of Ontario Institute of Technology, Canada. From 2015 to 2021, he was an Assistant Professor with HIT Shenzhen (HITSz). Since 2022, he has been an Associate Professor with the Peng Cheng Laboratory, Shenzhen. His research interests include satellite communications, resource allocation, and mobile Internet.



**Shaohua Wu** (Member, IEEE) received the Ph.D. degree in communication engineering from the Harbin Institute of Technology (HIT), Harbin, China, in 2009.

From 2009 to 2011, he held a post-doctoral research position with the Department of Electronics and Information Engineering, Shenzhen Graduate School, HIT Shenzhen, Shenzhen, China, where he was an Associate Professor from 2012 to 2020. Since 2021, he has been a Professor with the School of Electrical and Information Engineering, HIT Shenzhen, and also a Professor with the Peng Cheng Laboratory, Shenzhen. His current research interests include wireless image/video transmission, space communications, advanced channel coding techniques, and B5G wireless transmission technologies.



**Rongxing Lu** (Fellow, IEEE) received the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Waterloo, Canada, in 2012. He is currently the Mastercard IoT Research Chair, an University Research Scholar, and an Associate Professor with the Faculty of Computer Science (FCS), University of New Brunswick (UNB), Canada. Before that, he was an Assistant Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), Singapore, from April 2013 to

August 2016. He was a Post-Doctoral Fellow with the University of Waterloo from May 2012 to April 2013. He has published extensively in his areas of expertise and was the recipient of nine best (student) paper awards from some reputable journals and conferences. His research interests include applied cryptography, privacy enhancing technologies, and the IoT-big data security and privacy.

He was awarded the most prestigious “Governor General’s Gold Medal” and won the Eighth IEEE Communications Society (ComSoc) Asia-Pacific (AP) Outstanding Young Researcher Award in 2013. Currently, he serves as the Chair for IEEE ComSoc Communications and Information Security Technical Committee (CIS-TC) and the Founding Co-Chair for IEEE TEMS Blockchain and Distributed Ledgers Technologies Technical Committee (BDLT-TC). He is the Winner of (2016–2017) Excellence in Teaching Award, FCS, UNB.



**Qinyu Zhang** (Senior Member, IEEE) received the bachelor’s degree in communication engineering from the Harbin Institute of Technology (HIT) in 1994 and the Ph.D. degree in biomedical and electrical engineering from the University of Tokushima, Japan, in 2003.

From 1999 to 2003, he was an Assistant Professor with the University of Tokushima. He has been with the HIT Shenzhen (HITSz) since 2003, where he is currently a Full Professor and the Vice President. His research interests include aerospace communications and networks and wireless communications and networks. He has been awarded the National Science Fund for Distinguished Young Scholars, Young and Middle-Aged Leading Scientist of China, and the Chinese New Century Excellent Talents in University. He obtained three scientific and technological awards from governments.