

MSPA: Multislot Pilot Allocation Random Access Protocol for mMTC-Enabled IoT System

Jian Jiao¹, *Member, IEEE*, Liang Xu, Shaohua Wu², *Member, IEEE*, Rongxing Lu³, *Fellow, IEEE*,
and Qinyu Zhang⁴, *Senior Member, IEEE*

Abstract—To provide massive connectivity in massive machine-type communications (mMTCs) for the Internet of Things (IoT) system, a novel grant free random access protocol, called multislot pilot allocation (MSPA) is proposed in this article, where the user equipments (UEs) are permitted to jointly transmit randomly chosen pilot sequences along with their data packets over multislot to resolve intracell pilot collision. In addition, by utilizing the belief propagation tool for the MSPA protocol, the closed-form expressions to the access failure probability (AFP) and system throughput in a finite length regime are derived, which are highly desired for practical-interest mMTC network. Further, a guideline for certain mMTC scenarios that target urgent serving requirement UEs is also proposed to minimize the access latency and maximize the system throughput under diverse AFP constraints. Finally, the parametrical analysis of the MSPA protocol is given by theoretical proof and simulation verification, which shed light on the advantages of our MSPA protocol over the existing protocols in terms of achieving high throughput and shortening the access latency.

Index Terms—Finite-length analysis, grant-free random access, massive machine-type communications (mMTCs), multislot pilot allocation (MSPA) protocol.

I. INTRODUCTION

THE Internet of Things (IoT) has been regarded as an enabling technology for smart city, smart grid, and industrial automation, and has been extensively studied in both industry and academia [1]–[3]. Various applications for the IoT are foreseen to increase exponentially during the years ahead, which imposes critical challenges for the network to provide massive connectivity. As one of the key technologies in the IoT, massive machine-type communications (mMTCs) are promising to enable ubiquitous connectivity

Manuscript received January 17, 2021; revised April 8, 2021; accepted May 3, 2021. Date of publication May 18, 2021; date of current version December 7, 2021. This work was supported in part by the National Natural Sciences Foundation of China (NSFC) under Grant 62071141, Grant 61871147, Grant 61831008, and Grant 62027802; in part by the Natural Science Foundation of Guangdong Province under Grant 2020A1515010505; in part by the Guangdong Science and Technology Planning Project under Grant 2018B030322004; in part by the Guangdong Special Support Plan under Grant 2016TX03X226; and in part by the project The Verification Platform of Multitier Coverage Communication Network for Oceans under Grant LZC0020. (*Corresponding authors: Shaohua Wu; Qinyu Zhang.*)

Jian Jiao, Liang Xu, Shaohua Wu, and Qinyu Zhang are with the Communication Engineering Research Centre, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China, and also with Peng Cheng Laboratory, Shenzhen 518055, China (e-mail: jiaojian@hit.edu.cn; xuliang_hit@foxmail.com; hitwush@hit.edu.cn; zqy@hit.edu.cn).

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

Digital Object Identifier 10.1109/JIOT.2021.3081535

for low-power and low-rate user equipments (UEs) without human intervention by utilizing an efficient random access mechanism [4]–[6].

As is known that, in traditional cellular networks, each UE can be allocated a dedicated pilot sequence to access the system such that no intracell pilot collision occurs [7], [8]. However, this is impractical in future mMTC scenarios. Since the number of mMTC UEs is growing at an impressive rate and much larger than that of the available pilot sequences, pilot collision is unavoidable. Moreover, the intermittent access often leads to have an unpredictable, random subset of UEs being activated in a transmission frame [9]. Thus, dedicated pilot allocation in such case is infeasible, and decentralized assignment of pilot sequences and perform grant free access becomes a natural choice [7], [8].

A. Motivations and Related Works

In grant-free random access, each UE selects a pilot sequence at random from a predefined pool, and then sends in the uplink followed by the data information, which can alleviate the overload and pilot collision to some extent [10]. A significant amount of researches have been put forth in this research direction. In [11], a random access protocol is proposed by combining the properties of massive multiple-input-multiple-output (mMIMO), allowing for performing iterative belief propagation (BP) to decontaminate pilot signals and improving the throughput at the cost of excessive access delay. Björnson *et al.* [12] proposed a strongest user collision resolution (SUCR) random access protocol, where the system permits all active UEs to select pilot sequences at random in an attempt to access. Then, the UE with the strongest channel gain is regarded as the contention winner among the UEs who choose the same pilot, which enables to resolve the pilot collisions in a distributive manner. However, as pointed out in [13], the SUCR protocol is unfair to some extent for UEs with weak channel gain. As an extension, a SUCR combined idle pilots access (SUCR-IPA) protocol is proposed in [14], where the UEs with weak channel gain are also permitted to contend for the pilots that are not selected by any UE, called idle pilots. This implies that the weaker UEs can try more attempts to access pilots, allowing for compensating on the UE fairness and increasing system throughput. However, in a crowded scenario, since the number of idle pilots decreases dramatically, there is no significant performance improvement

on throughput and access failure probability (AFP), i.e., the probability that a UE cannot be resolved by the base station (BS), and the pilot contamination becomes the bottleneck of AFP [15].

Note that the random access to pilots with the tradeoff between the AFP and the access latency is a new shift in the design of flexible pilot assignment protocols for mMTC networks, which however is not explicitly considered in the existing works. In [10], a coded pilot access protocol for crowded scenarios is proposed, where the pilot allocation is considered in several time slots when the data part is repeated, which enables the employing of the bipartite graph to cancel the interference of copies in multiple pilot selections. In [16], a super preamble consisting of multiple consecutive preambles is proposed for grant-free random access when the active UEs are less than 20, where the UE randomly selects a preamble sequence in each preamble phase, and transmits a super preamble followed by a data payload to the BS. This allows for reducing the round-trip signal overhead between the BS and UEs [17]. Moreover, consider that the typical mMTC can be characterized by delay tolerant and medium reliability, there exists an inherent tradeoff between the minimum access time slots and the achievable AFP under the limited pilot sequences, which allows for guaranteeing the desired AFP of the UEs at the price of slightly increasing the access time slots. Thus, in this article, we propose a multislot pilot allocation (MSPA) random access protocol, investigate the performance of the MSPA random access protocol in the mMTC scenarios that target on an optimal number of access time slots under required AFP thresholds, and elaborate on the superiorities over the existing random access protocols in overload case.

The AND-OR tree is acknowledged for evaluating the performance of the random access protocol, and has shown a satisfying performance [13], [14], [18]. This is mainly due to that a set of contaminated pilot signals can be viewed as a graph code, and the iterative cancellation process is exactly the same as graph-based codes, e.g., LT and LDPC codes [19], [20], allowing for the exploiting AND-OR tree to track the evolution of the AFP. However, the inaccuracy of the asymptotic AND-OR tree result in a finite regime, has also been noted in previous works [21]–[23]. It is easy to understand that for a crowded scenario with a large number of UEs and limited pilots, the iteration process in AND-OR tree can be stopped either by the nonempty stopping set or reaching the maximum access time. Up to now, there have been a few investigations into the finite analysis. In [21] and [24], tight analytical approximations to the packet loss probability of irregular repetition slotted ALOHA protocols at a low channel load regime are performed. Further, Lazaro and Stefanović [23] proposed an exact finite-length analysis of frameless ALOHA with performance optimization. However, such a finite-length analysis on predicting the performance of pilot allocation random access schemes is still missing, which calls for the finite-length analysis of random access protocols under more realistic assumptions. Therefore, in this article, we investigate a finite-length analysis framework to accurately

estimate the AFP of the proposed MSPA scheme in a finite length regime.

B. Contributions

Specifically, the main contributions of this article are three-fold as follows.

- 1) *MSPA Scheme*: We propose an MSPA random access protocol design for the massive mMTC system, where the UEs jointly transmit randomly chosen pilot sequences along with their data packets in multislot, allowing for introducing the graph representation of the MSPA protocol by denoting the variable nodes (VNs) as the active UEs and factor nodes as the selected pilots in certain access time slots. We then illustrate how to resolve the UEs via a successive interference cancellation (SIC) procedure. We show that in the case of overload, our MSPA scheme can significantly outperform the relevant random access schemes in terms of achieving lower AFP and higher throughput. That enables us to highlight the advantages of our MSPA scheme which aims to provide massive connectivity in the mMTC system.
- 2) *Finite-Length Analysis and Guide Line for MSPA Scheme*: Utilizing the analogy between the iterative SIC process of the MSPA protocol and the BP decoding for an erasure code over bipartite graph [25], we derive the closed-form expressions to the AFP and system throughput in a finite length regime, which indeed exists in the most practical mMTC systems. Further, we show how the derived expressions can be used to design a satisfying MSPA system in which some UEs require to be served with desired AFP, i.e., target on high access reliability. Accordingly, a guideline is proposed for the MSPA scheme by jointly optimizing the allocated pilots and minimum access latency under diverse AFP constraints.
- 3) *Parametrical Analysis and Simulation Results*: For the sake of complete analysis, we present parametrical analyses for our MSPA random access scheme by theoretical proof and simulation results, which illustrate how the access latency affects the system load and the system throughput under different AFP constraints. Thus, the MSPA random access protocol can significantly improve both the AFP and system throughput when extending the access latency properly. Further, we present the simulation implementation of our derived optimal transmission policy that highlights its performance and showcase interesting insights on the differences between our MSPA random access protocol and several benchmark protocols.

The remainder of this article is organized as follows. The system model and the principles of the MSPA protocol are given in Section II. Section III elaborates the finite-length analysis of the proposed MSPA protocol. Section IV depicts the parametrical analysis of the MSPA protocol. After that, the simulation results and conclusion are presented in Sections V and VI, respectively.

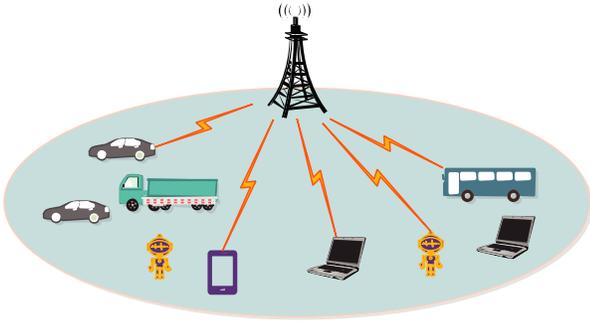


Fig. 1. Single-cell crowded scenario under our consideration.

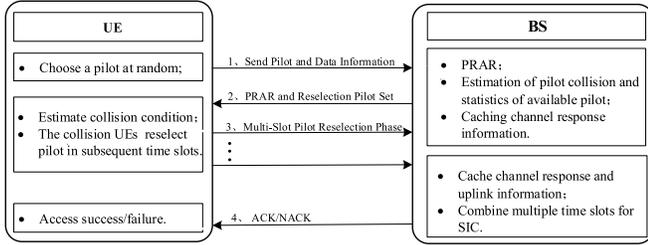


Fig. 2. MSPA protocol.

II. SYSTEM MODEL AND THE PROPOSED MSPA PROTOCOL

A. System Model

We consider a cellular network in which each BS is equipped with M antennas, and there are K active UEs each with a single antenna and a BS with M antennas, as shown in Fig. 1, and the transmission is organized in blocks of consecutive time slots, referred to a frame, where the time–frequency resource is divided into two blocks, namely, the pilot random-access block and the payload data block. Specifically, the payload data block is used for transmitting UEs' data packet, and the pilot random-access block is utilized for accessing the pilot [26].

In the conventional grant-based communication system, the dedicated pilot allocation has shown a satisfying performance, as each UE can be preallocated the pilot sequence to complete data transmission with no pilot collision. Then, the BS can estimate the channel responses between the UEs and BS according to their dedicated pilots, and perform decoding in a computationally efficient manner. However, such a dedicated pilot allocation is impractical in future mMTC networks where the number of UEs is far much larger. To address this issue, we resort to the MSPA random-access protocol.

B. Proposed MSPA Protocol

Fig. 2 shows the four steps of the MSPA random-access protocol. Let τ_p denote the number of available pilots allocated to K active UEs, denoted by $\mathbf{U} = \{U_1, U_2, \dots, U_K\}$. We assume that each frame contains ΔN time slots, and we divide each frame into two stages, i.e., the first stage with one time slot and the second stage with the remaining $\Delta N - 1$ time slots. Without loss of generality, each duration of the time slot

is equal and normalized as 1. Thus, the access latency of our MSPA random-access protocol is ΔN .

Prior to step 1, we assume that each UE will receive a control signal broadcasted by the BS, allowing for estimating the average channel gain and synchronizing with the BS. We define ρ_k as the uplink transmitting power of U_k , which is inversely proportional to the large-scale fading, denoted by γ_k . Then, $\rho_k = (\bar{\rho})/(\gamma_k)$, where $\bar{\rho}$ is the average received power at the BS as considered in [27]. Let h_k denote the channel gain between U_k and the BS, where $h_k = \sqrt{\gamma_k}h'_k \in \mathbb{C}^M$, and h'_k is the small-scale fading.

Note that the most common traffic in mission critical communications (MCCs) and uplink mMTC applications are short packet communications, which only involves several hundreds of bits and the data can fit in a single payload packet. Grant free access is adopted in our MSPA random-access scheme since it reduces the signaling overhead caused by conventional grant-based schemes. The details of this protocol are described as follows.

Step 1: The K UEs randomly select a pilot from the mutually orthogonal pilot set $\mathbf{D} = \{\xi_1, \xi_2, \dots, \xi_{\tau_p}\}$ and send it to the BS together with the uplink information. The uplink information includes the selected pilot index (i.e., the header of the data packet) and the encoded data (i.e., the payload of the data packet). Each UE selects a certain pilot with equal probability $1/\tau_p$. The pilot collision will occur if more than one UE selects the same pilot.

The pilot signals received at the BS, denoted by $\mathbf{Y}_q^p \in \mathbb{C}^{M \times \tau_p}$, are as follows [12], [28]:

$$\mathbf{Y}_q^p = \sum_{k \in \mathbf{U}} \sqrt{\rho_k} h_k^q (\xi_k^q)^T + \mathbf{N}_p \quad (1)$$

where h_k^q and ξ_k^q are the channel gain between U_k and the BS and the selected pilot ξ_j in the q th time slot, respectively, and $\mathbf{N}_p \in \mathbb{C}^{M \times \tau_p}$ is the independent receiver noise with each entry being a circularly symmetric complex Gaussian random variable with mean zero and variance θ^2 , i.e., $CN(0, \theta^2)$.

Based on the received pilot signal, the BS estimates the channel response corresponding to pilot sequence ξ_j , denoted by Ω_q^j , and can be expressed as follows [13], [29]:

$$\Omega_q^j = \mathbf{Y}_q^p \frac{\xi_j^*}{\|\xi_j\|} = \sum_{k \in Z_q^j} h_k^q + \mathbf{N}_p \frac{\xi_j^*}{\|\xi_j\|} \quad (2)$$

where Z_q^j denotes the set of UEs that select the pilot ξ_j in the q th time slot, “*” denotes the complex conjugate of a vector, and $\|\cdot\|$ denotes the Euclidean norm of a vector.

Step 2: The BS caches the channel response estimated in step 1 and obtains the precoded random-access response (PRAR) based on (2), denoted by \mathbf{V}_q , and we have

$$\mathbf{V}_q = \sum_{j=1}^{\tau_p} \frac{\Omega_q^j}{\|\Omega_q^j\|} (\xi_j)^T. \quad (3)$$

Let $\hat{\alpha}_j$ denote the sum of gains and interference corresponding to ξ_j at the BS. When the number of antennas M is asymptotic large, $\hat{\alpha}_j$ can be estimated by the BS as

follows [16]:

$$\hat{\alpha}_j = \frac{\|\Omega_q^j\|^2}{M} = \sum_{i \in Z_q^j} \rho_i \gamma_i \tau_p + \theta^2 = \bar{\rho} \tau_p |Z_q^j| + \theta^2. \quad (4)$$

Then, the BS can identify the state of pilot sequence in the q th slot through the value of $\hat{\alpha}_j$ as follows:

$$|Z_q^j| = \frac{\hat{\alpha}_j - \theta^2}{\bar{\rho} \tau_p}. \quad (5)$$

Thus, the state of pilot sequences can be divided into the following three cases.

Case 1: $|Z_q^j| = 0$ indicates that no UE selects the pilot and the pilot is the idle pilot.

Case 2: $|Z_q^j| = 1$ indicates that only one UE selects the pilot and these pilots are called the singleton pilot.

Case 3: $|Z_q^j| > 1$ indicates that multiple UEs select the pilot simultaneously, called the collision pilot.

The idle pilots and collision pilots constitute the reselection pilot set, denoted by \mathbf{R} . Then, the BS can estimate the number and index of remaining pilots in \mathbf{R} [30], as well as the number of UEs involved in the collision in step 1.

After these processes, the BS broadcasts PRAR information and the reselection pilot set \mathbf{R} over the downlink channel to all active UEs. Note that the BS estimates the channel response to the received pilot sequence at each access time slot for the later pilot collision identifying and status packets decoding as the same processes without feedback to the UEs.

Step 3: By detecting PRAR information received from the BS, each UE can independently determine whether they have generated a collision at step 1. The UEs that select the singleton pilot, called singleton UEs, no longer send new pilot and data messages, while the UEs that select the collision pilot, called collision UEs, randomly reselect the pilot from set \mathbf{R} in subsequent time slots. Similarly, the BS performs pilot identification and caches the state of pilots and its status packets in each access time slot for the later decoding. At the receiving end, the BS uses the received pilot signals for channel estimation, and caches the estimated channel response and uplink data information for subsequent processing.

Step 4: By utilizing the information cached in the previous steps, the BS performs SIC to recover UEs' information. The BS can locate the UEs which can be successfully decoded immediately, that is, no collision occurs, i.e., $|Z_q^j| = 1$. Once a UE is resolved, the selected slot index and the pilot will be located to eliminate the interference from the copies of the recovered UEs. This allows the potential of having more singleton pilots and continuing the process of iterative elimination. If the decoding is successful, the BS sends ACK to the recovered UE. Otherwise, it sends a NACK.

C. SIC Algorithm of the MSPA Protocol

As the BS does not have *a priori* knowledge of the random activity and pilot choices of the K UEs, the pilot contamination cannot be solved by using Gaussian elimination. Instead, we employ SIC, as in recent works on pilot allocation random

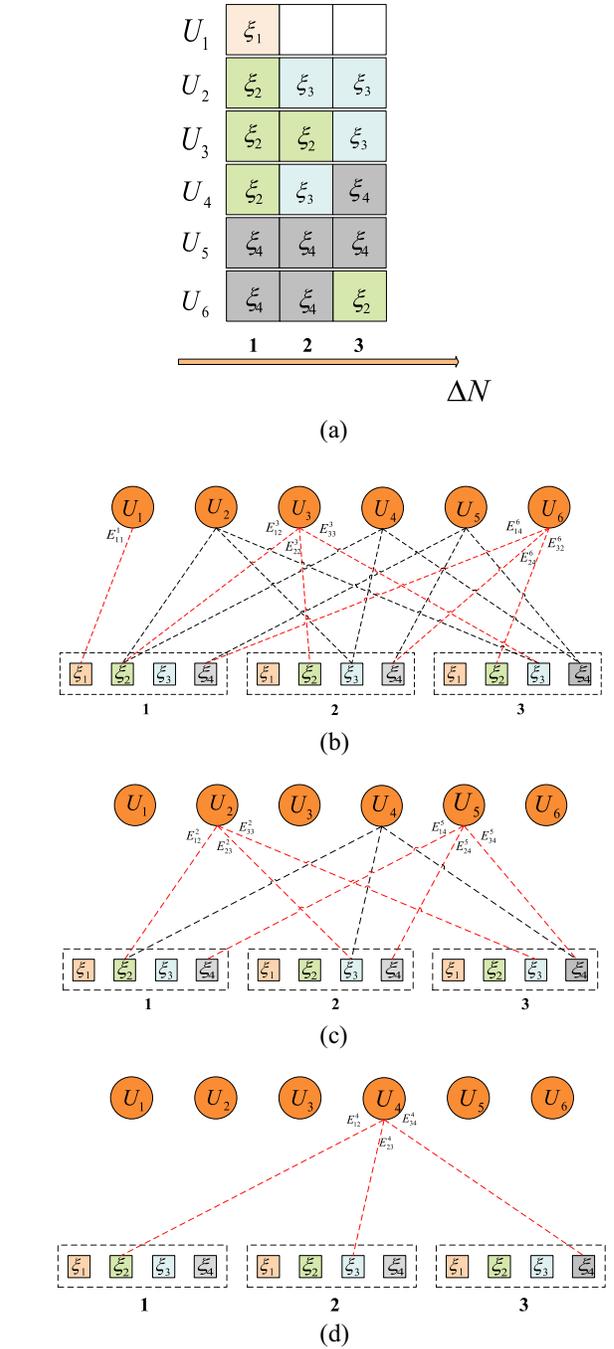


Fig. 3. Bipartite graph representation of the MSPA scheme and an example of the SIC process. (a) Example of the pilot schedule in a frame with $\Delta N = 3$, $K = 6$ UEs, and $\tau_p = 4$ pilots. (b) Bipartite graph representation of the example depicted in (a). (c) Second iteration of SIC. (d) Third iteration of SIC.

access [31]–[33]. The SIC applied in our context is presented as follows.

In a practical random-access system, there may exist a potential risk of being undecodable at the BS due to the accumulation of noise during interference cancellation. In our considered scheme, we assume that noise does not garble the decoding process, that is, when a pilot has been reduced to degree one, the corresponding UE is resolved successfully. Initially, the BS immediately locates the singleton pilots,

i.e., case 2 of $|Z_q^j| = 1$, and decodes the UE. Further, the interference caused by the associated pilot transmissions can be subtracted from the superposition of the UE channels, allowing for the new case 2 of $|Z_q^j| = 1$, whereby more UEs can be resolved, and the iterative cancellation process can continue. As the employed decoding algorithm is analogous to BP decoding of erasure codes, a bipartite graph considered as a common way to visualizing such codes is constructed to describe the MSPA protocol. Consider the bipartite graph representation of MSPA, where each of the K UEs is represented by a VN, and the τ_p pilots allocated in each time slot is represented by a check node (CN), we present a bipartite graph in Fig. 3(b), which corresponds to the example shown in Fig. 3(a).

Example: To facilitate the analysis, we denote E_{ij}^k as the edge that connects the UE U_k and the pilot ξ_j in the i th time slot. As shown in Fig. 3(b), the BS detects according to the header of the data packet that, in the first time slot, E_{11}^1 connects to the singleton pilot ξ_1 , and thereby directly recovers U_1 . Similarly, U_3 and U_6 can also be recovered by performing such a detection for singleton pilots. As a result, E_{12}^3 , E_{22}^3 , E_{14}^6 , and E_{24}^6 will be deleted after the first iteration, which enables the cancellation of its interference in the following iterations. As illustrated in Fig. 3(c), the cancellation of interference from U_3 and U_6 allows for the resolution of U_2 and U_5 . As U_2 and U_5 connect to the singleton pilot ξ_3 and ξ_4 through E_{33}^2 and E_{14}^5 , respectively, which can be detected by the BS in the second iteration. During the third iteration in Fig. 3(d), U_4 can be resolved and thus all of the UEs are resolved with the help of our MSPA random-access protocol. In the i th iteration of SIC, we define ψ_i^{jq} as the channel response corresponding to pilot ξ_j in the q th time slot and the i th iteration, Φ_i^j as the index of the UE whose uplink message can be decoded correctly by utilizing ψ_i^{jq} , $\tilde{h}_{\Phi_i^j} = \{\tilde{h}_{\Phi_i^j}^1, \tilde{h}_{\Phi_i^j}^2, \dots, \tilde{h}_{\Phi_i^j}^M\}$ as the estimated channel response of UE Φ_i^j , A_i as the set of indexes of pilots embedded in the successfully decoded uplink message. Let H denote the recovered channel responses via the proposed SIC algorithm. Detailed iterations of the SIC algorithm are presented in Algorithm 1.

Note that M has an impact on the asymptotic favorable propagation of the mMIMO system, and degrades the accuracy of detecting the types of pilot especially when M takes a small value. As illustrated in the SIC algorithm of the MSPA random-access scheme, the BS should first locate those degree 1 factor nodes in the bipartite graph to start the SIC process, and has the knowledge that the pilots selected by individual UEs to obtain a new bipartite graph. Then, we focus on the reliability of detecting singleton pilots with different values of M . There are two types of error that will influence the accuracy of detecting singleton pilots. One is that whether an idle pilot is detected to be the singleton pilot. The other is that whether a collision pilot is detected to be the singleton pilot. We define P_{idle} as the probability that an idle pilot is detected correctly, instead of being detected as a singleton pilot or collision pilot.

Fig. 4 shows P_{idle} versus M , where SNRe is the SNR at the corner of the cell. We can see that P_{idle} is almost equal to 1 for SNRe = 6 and 9 dB when $M > 50$. Note that an idle pilot

Algorithm 1: SIC Algorithm of the MSPA Protocol

Input: Y_n^p, Ω_q^j, V_q^j

- 1: Set $|A_0| = 1, H = \emptyset, i = 1, \psi_1^{jq} = \Omega_q^j$
- 2: **while** $|A_{i-1}| \neq 0$ **do**
- 3: $A_i = \emptyset$
- 4: **for all** $j \in \{1, 2, \dots, \tau_p\}$ and $q \in \{1, 2, \dots, \Delta N\}$ **do**
- 5: Use ψ_i^{jq} ($j \in A_{i-1}$) to detect uplink data
- 6: **if** the uplink message can be decoded correctly **then**
- 7: The BS detects that the pilot $\chi_{q'j}^1$ and $\chi_{q''j}^2$ are connected with UE Φ_i^j , and $\tilde{h}_{\Phi_i^j} = \psi_i^{jq}$ ($j \in A_{i-1}$)
- 8: Update the channel response: $\psi_{i+1}^{\chi_{q'j}^1} = \psi_i^{\chi_{q'j}^1} - \tilde{h}_{\Phi_i^j}$ and $\psi_{i+1}^{\chi_{q''j}^2} = \psi_i^{\chi_{q''j}^2} - \tilde{h}_{\Phi_i^j}$.
- 9: Update the pilot index set and the channel response set: $A_i = A_i \cup \{\chi_j^{1q'}, \chi_j^{2q'}\}$ and $H = H \cup \tilde{h}_{\Phi_i^j}$
- 10: **end if**
- 11: **end for**
- 12: **for all** $m \in \{1, 2, \dots, \tau_p\}$ and $q \in \{1, 2, \dots, \Delta N\}$ **do**
- 13: Update the channel response: $\psi_{i+1}^{mq} = \psi_i^{mq}$
- 14: **end for**
- 15: $i = i + 1$
- 16: **end while**

Output: H : the set of the recovered channel response

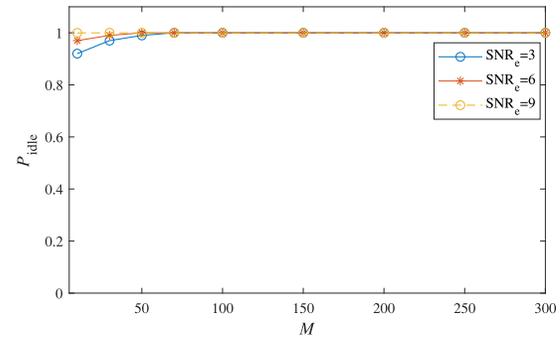


Fig. 4. Probability P_{idle} versus M .

is unlikely to detect to be a singleton pilot as the value of M in the practical mMIMO system is much larger than 100.

Moreover, we can observe from (2) and (3) that the estimated channel response corresponding to a collision pilot is a sum of the channel responses of all the contenders selecting the same pilot. Then, if the collided pilots are detected to be singleton pilots, the BS will definitely fail to decode the corresponding the uplink message as the decoded uplink message by a collision pilot is a summation of uplink message of the contenders [13]. Note that a collision pilot that is detected to the singleton pilot has no impact on the system throughput in our MSPA scheme. Therefore, the complexity of the SIC in our MSPA scheme is $O((\tau_p \delta N)^2)$, since the received pilot signals have $\tau_p \delta N$ dimensions at most.

D. Total Interference Power in the MSPA Protocol

When a collided pilot is reduced to a singleton pilot in the bipartite graph, we prove that our SIC algorithm can resolve the uplink data corresponding to the singleton pilot with high reliability. In the following, we first derive the expression of the total number of collided UEs in a time slot. Then, we derive the expression of the total interference power in the uplink random-access system.

Considering each UE activates independently by a certain probability in each time slot, we assume that the number of active UEs in a certain time slot satisfies a Poisson distribution, and we define μ as the average number of active UEs for per time slot. The probability of a UE fails in a collision, denoted by f , can be expressed as $1 - \exp(-\mu/\tau_p) = f$. Therefore, the average number of UEs in a same pilot follows the Poisson distribution with μ/τ_p .

Let E_s denote the event that there are s singleton pilots when $\tau_p = E$, i.e., there are s singleton UEs and each choose one of these pilots. We define as the event C_a that in the collision pilots, the total number of collision UEs is a . The corresponding mathematical expressions can be expressed by the piecewise functions as follows:

$$\begin{cases} \Pr(C_0|E_s) = \left(\frac{\exp(-\frac{\mu}{E})}{\Pr(E_1)}\right)^{E-s} \\ \Pr(C_1|E_s) = 0 \\ \Pr(C_a|E_s) = \frac{\Pr_s(C_a) - \Pr(C_a, \bar{E}_s)}{\Pr(E_{E-s})}, \quad a \geq 2, s \leq E-2 \\ \Pr(C_a|E_s) = \frac{(\frac{\mu}{E})^a \exp(-\frac{\mu}{E})}{a! \Pr(E_1)}, \quad a \geq 2, s = E-1 \end{cases} \quad (6)$$

where $\Pr(E_s)$, $\Pr_s(C_a)$, and $\Pr(C_a, E_s)$ can be expressed as

$$\begin{cases} \Pr(E_s) = \left(1 - \frac{\mu}{E} \exp(-\frac{\mu}{E})\right)^s, \\ \Pr_s(C_a) = \frac{((L-s)\frac{\mu}{E})^a \exp(-\frac{\mu}{E})}{a!} \\ \Pr(C_a, \bar{E}_s) = \sum_{i=1}^{\sigma} \binom{s}{i} \Pr(\bar{E}_1)^i \Pr(C_{a-i}, E_{s-i}) \end{cases} \quad (7)$$

where $\sigma = \min\{E-s, a\}$, $\Pr(\bar{E}_1) = (\mu/E) \exp(-\mu/E)$, and $\Pr(C_{a-i}, E_{s-i}) = \Pr(C_{a-i}|E_{s-i}) \Pr(E_{s-i})$.

Note that the total interference will hamper the data decoding. We use the total number of collided UEs and the distribution of received power to calculate the total interference power. Assume that the transmission power of all UEs is P_T . Hence, under the Rayleigh fading model, the distribution of received power, denoted by $f_{sR}(p)$, can be expressed as [34]

$$f_{sR}(p) = \frac{1}{2b} \exp\left(-\frac{p}{2bP_T}\right) \quad (8)$$

where $2b$ is the average power of the multipath component. We select the parameter in [34] as $b = 0.251$. Then, we can derive the distribution of total interference power, denoted by $g(I|E_s)$, as follows:

$$g(I|E_s) = \sum_{n=0}^{\infty} f_{sR}(I) \otimes_{n-1} f_{sR}(I) \times \Pr(C_n|E_s) \quad (9)$$

where \otimes_{n-1} denotes $n-1$ times convolution.

Fig. 5(a) shows the collision probability versus the total number of collision UEs a for different values of f . We can see that the collision probability is fluctuating over the

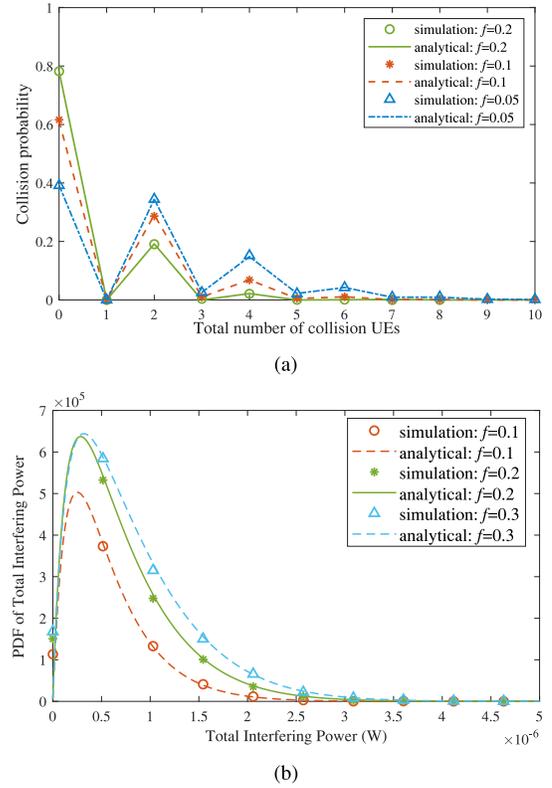


Fig. 5. (a) Collision probability and (b) total interference power under the Rayleigh fading.

number of collision UEs. Specifically, the collision probability is lower when a is odd, and is higher when a is even. It is mainly because the probability of three UEs chooses the same collision pilot [i.e., $\Pr(C_3|E_{s-1})$] is lower than the probability of two UEs choose the same collision pilot [i.e., $\Pr(C_2|E_{s-1})$]. Moreover, for the case of $a = 4$ [i.e., $2\Pr(C_2|E_{s-2}) + \Pr(C_4|E_{s-1})$], it means that four UEs choose two same collision pilots (each of the collision pilot is selected by two UEs) or four UEs choose the same pilot, which is more possible than the case of $a = 3$. Moreover, the smaller f , the smaller the probability that a UE has a collision with other UEs, the collision probability is decreasing with the increase of f when a takes a small value.

Fig. 5(b) shows the PDF of total interference power for different values of f . We can see that the total interference power point where the probability density of total interference power achieves the maximum will increase with the increase of f . The total interference power is below a quite smaller value 10^{-6} W. Therefore, it is reasonable in this article to assume that noise is not garbled the decoding process of singleton UEs.

III. FINITE-LENGTH ANALYSIS

In this section, we mainly analyze the performance of the considered MSPA scheme, including the AFP, i.e., the probability that a UE's packet is recovered by the BS, and the throughput of the system. By utilizing the finite length scaling analysis of erasure codes over the bipartite graph [25], [35], we derive the expressions to the AFP in the finite length regime.

Further, we show how the derived expressions can be used to design system that can guarantee diverse AFP requirements with acceptable access latency for the MSPA scheme, that is, minimizing the access latency with AFP below a required threshold.

A. Degree Distributions

Let g denote the ratio of the number of active UEs to that of available pilots, namely, system load. According to the protocol depicted in Fig. 2, in the first time slot, each active UE selects a pilot at random from a predefined pool, and then sends in the uplink followed by the data information. Note that the pilot assignment is randomized in each time slot when the data part is repeated. Thus, the pilot selection for any active UE has the following cases.

- 1) The UE selects the singleton pilot and remains silent in subsequent $\Delta N - 1$ time slots instead of the reselecting pilot.
- 2) The UE selects a collision pilot and reselects a pilot from \mathbf{R} in subsequent $\Delta N - 1$ time slots.

Let Λ_l denote the probability that the degree of a UE is l and P_{nc} denote the probability that the UE selects a singleton pilot in the first time slot, which can be calculated by the following:

$$P_{nc} = \left(1 - \frac{1}{\tau_p}\right)^{K-1}. \quad (10)$$

The probability that a pilot collision occurs is

$$P_c = 1 - \left(1 - \frac{1}{\tau_p}\right)^K - \left(1 - \frac{1}{\tau_p}\right)^{K-1} \quad (11)$$

where $(1 - 1/\tau_p)^K$ is the probability that none of the active K UEs selects the pilot and $(1 - 1/\tau_p)^{K-1}$ that there is only one UE selects the pilot.

In our MSPA protocol, the pilot reselection takes place in multiple time slots such that the failed UE in the first time slot can try more access attempts in the remaining $\Delta N - 1$ time slots. We define ε as the probability that a UE cannot be resolved by the SIC at the end of the transmission frame.

In the case of pilot reselection, the MSPA protocol admits all of the collision UEs to reselect pilot from the set \mathbf{R} in subsequent $\Delta N - 1$ time slots, which means that the degree of each UE is 1 or ΔN . Thus, Λ_l can be expressed as

$$\Lambda_l = \begin{cases} \left(1 - \frac{1}{\tau_p}\right)^{K-1}, & l = 1 \\ 1 - \left(1 - \frac{1}{\tau_p}\right)^{K-1}, & l = \Delta N. \end{cases} \quad (12)$$

The polynomial representations can also be obtained by

$$\Lambda(x) = \sum_d \Lambda_d x^d \quad (13)$$

where x denotes the variable of the polynomial representations.

Let Ω_i^j denote the probability that a pilot is selected by l UEs in the i th time slot, which can be calculated by the following:

$$\Omega_i^j = \begin{cases} \binom{K}{l} \left(\frac{1}{\tau_p}\right)^l \left(1 - \frac{1}{\tau_p}\right)^{K-l}, & i = 1 \\ \binom{F}{l} \left(\frac{1}{R}\right)^l \left(1 - \frac{1}{R}\right)^{F-l}, & i > 1 \end{cases} \quad (14)$$

where $R = |\mathbf{R}|$ is the expected number of reselection pilots, and F is the expected number of UEs that will reselect pilots from \mathbf{R} in step 3.

B. AFP and Throughput Analysis

In the following, we first derive the expression of ε by utilizing the bipartite graph analysis [15], [22], and then obtain the expressions of AFP and the system throughput for the considered system. For each UE, since the pilot selection is randomized in different time slots, and follows the VN degree distribution $\Lambda(x)$, the corresponding bipartite graph has edge-perspective VN degree distribution as follows:

$$\lambda(x) = \sum_d \lambda_d x^{d-1} \quad (15)$$

where $\lambda_d = [(d\Lambda_d)/(\sum_d d\Lambda_d)]$.

If the number of time slots and pilots tend to be a large value in the practical mMTC scenario, the number of access attempts in a slot follows a Poisson distribution, and the edge-perspective CN degree distribution, denoted by $\Psi(x)$, can be obtained by

$$\Psi(x) = e^{-\frac{K\Lambda'(1)(1-x)}{\tau_p\Delta N}} \quad (16)$$

where $\Lambda'(x)$ denotes the derivative of $\Lambda(x)$.

We define $G = [K/(\tau_p\Delta N)]$ as the average number of UEs that select a pilot per time slot. Note that as K tends to be a large value, thus there exists a BP decoding threshold as G^* , that is, ε tends to zero for $G < G^*$ and to one for $G > G^*$.

As the SIC operations is similar with the BP decoding for erasure code over the same bipartite graph as shown in Fig. 3, where the CNs represent the codewords. Thus, we can model the UEs behavior in the pilot selection and the SIC procedure in the finite length regime. In [25], a generalized finite-length analysis of the erasure code is proposed, where the block error probability as a function of erasure probability δ is approximated as [35]

$$P_{\text{BEP}} = Q\left(\frac{\sqrt{n}(\delta_r^* - \beta_r n^{-2/3} - \delta)}{\alpha}\right) + O(n^{-1/3}) \quad (17)$$

where $\alpha = \sqrt{\alpha_r^2 + \delta(1-\delta)}$, δ_r^* is the BP decoding threshold of the erasure code, n is the length of the codewords, and r is the nominal rate. β_r and α_r are constants that depend on r , and $Q(x)$ is the tail probability of the standard normal distribution.

Thus, we denote m as the input information corresponding to the codewords n . By utilizing the mapping relationship of BP decoding between our MSPA random-access protocol and the erasure code, G can be written as a function of δ and m as follows:

$$G = \frac{K}{\Delta N \tau_p} = \frac{\delta n}{m}. \quad (18)$$

Moreover, by substituting the BP decoding threshold δ^* into (18), we can obtain that

$$G^* = \frac{\delta^* n}{m} = \frac{\delta^*}{1-r} \quad (19)$$

where $r = (n-m)/n$.

TABLE I
PARAMETERS OF (20) FOR THE MSPA RANDOM-ACCESS PROTOCOL

ΔN	α_0	β_0	G^*	σ
2	0.47785	0.97684	0.46727	0.67863
3	0.45762	0.91247	0.41014	0.74386
4	0.41874	0.86646	0.36245	0.84201
5	0.38942	0.78114	0.32478	0.92156
6	0.34781	0.71417	0.29984	0.96454

Let g denote the system load, i.e., $g = K/\tau_p$. Using $G = [K/(\tau_p \Delta N)]$ together with (19) substituting in (17), we can obtain the expression to ε in the following lemma.

Lemma 1: For the MSPA scheme, the probability that a collision UE remains unresolved in the pilot resection is given as

$$\varepsilon \approx Q \left(\frac{\sqrt{\Delta N \tau_p} (G^* - \beta_0 (\Delta N \tau_p)^{-2/3} - g/\Delta N)}{\sqrt{\alpha_0^2 + g/\Delta N}} \right)$$

where α_0 and β_0 are scaling parameters for $r = 0$, i.e., $m = n$, and computed by the method introduced in [36], and also shown in Table I.

The proof of this lemma and the derivation of α_0 and β_0 will be given in Appendix A.

As the MSPA protocol is operated with multislot with limit of available pilots, the ε in Lemma 1 agrees well with the simulation result in the low-to-medium system load cases ($0.2 \leq g \leq 1.4$), but deviates in the high system load cases [35]. Therefore, we suggest that in the overload scenario, a suitable correcting parameter, denoted by σ , shall be considered in the prediction of ε . We can update the expression of ε as follows:

$$\varepsilon = \sigma Q \left(\frac{\sqrt{\Delta N \tau_p} (G^* - \beta_0 (\Delta N \tau_p)^{-2/3} - g/\Delta N)}{\sqrt{\alpha_0^2 + g/\Delta N}} \right) \quad (20)$$

where the correction coefficient σ is the value of ε for the case of $g \rightarrow 1$, which is computed by the density evolution [36], as shown in Table I.

We define P_M as the average AFP of the MSPA protocol, i.e., the probability that a UE fails to access the system in one-time transmission, and T denotes the throughput of our MSPA protocol, i.e., the average sum of resolved UEs per time slot in one-time transmission. By utilizing the degree distributions of VNs in (12) and expression of ε , the following lemma is then derived.

Lemma 2: For the MSPA scheme, P_M and T are given as

$$P_M = \varepsilon \Lambda_{\Delta N} \quad (21)$$

and

$$T = \frac{K(1 - P_M)}{\Delta N}. \quad (22)$$

IV. OPTIMAL ANALYSIS OF THE MSPA PROTOCOL

In this section, we employ the finite-length analysis to analyze the parametric performance of the MSPA protocol,

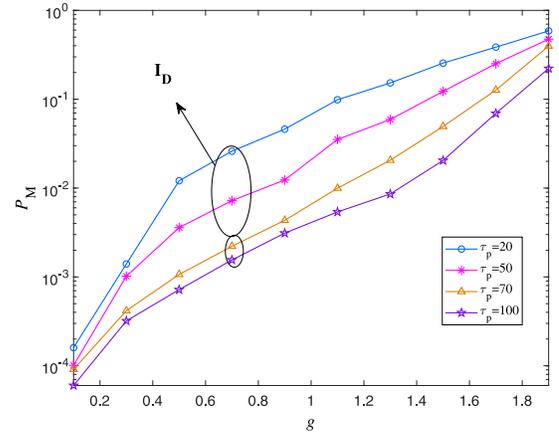


Fig. 6. Impact of τ_p on P_M , where $\tau_p^- = 100$ and $10 \leq K \leq 250$.

and formulate an optimization problem, aiming to minimize the access latency and the number of required pilots by given diverse AFP and system load constraints. Moreover, the maximum throughput is also analyzed in our work.

A. Minimization of the Access Latency

It is straightforward that assigning parameters τ_p and ΔN enables the system to guarantee the performance. This is due to the fact that the AFP of the MSPA protocol decreases with the increase of ΔN , and is sensitive to the variations in τ_p . When the value of ΔN and the AFP threshold are given, there exists a minimum τ_p , denoted by τ_p^* , that satisfies the following conditions.

- 1) If $\tau_p \geq \tau_p^*$, then $P_M \leq P_M^*$.
- 2) If $\tau_p < \tau_p^*$, then $P_M > P_M^*$.

We present the following proposition with the analysis on how τ_p is possible to improve the AFP, and elaborate the sensitivity of AFP on τ_p .

Proposition 1: Let τ_p^- denote the maximum number of available pilots in the system. There exists an interval for τ where P_M decreases rapidly with the increasing of τ_p , denoted by I_D and given as $[I_D = (0, \min(\tau_p^-, g\tau_p^-))]$.

Proof: The proof is given in Appendix B. ■

Fig. 6 shows the impact of τ_p on P_M . We can observe that for $g = 0.7$, P_M rapidly increases with the increasing of τ_p in the interval of I_D , while it increases slightly with the increase of τ_p out of the interval of I_D , i.e., P_M increases dramatically from $\tau_p = 20$ to $\tau_p = 70$, and increases at a slower pace when $\tau_p > 70$. This implies that P_M is very sensitive to the variations in τ_p in I_D . Similarly, when the value of τ and the AFP threshold are given, there exists a minimum ΔN , denoted by ΔN^* , that satisfies the following conditions.

- 1) If $\Delta N \geq \Delta N^*$, then $P_M \leq P_M^*$.
- 2) If $\Delta N < \Delta N^*$, then $P_M > P_M^*$.

As discussed in (12) and (20), we have $\Lambda_{\Delta N} = 1 - (1 - 1/\tau_p)^{K-1}$ and

$$\varepsilon = \sigma Q \left(\frac{\sqrt{\Delta N \tau_p} (G^* - \beta_0 (\Delta N \tau_p)^{-2/3} - g/\Delta N)}{\sqrt{\alpha_0^2 + g/\Delta N}} \right).$$

By substituting (12) and (20) into (21), we have

$$P_M(\tau_p, \Delta N) = \left[1 - \left(1 - \frac{1}{\tau_p} \right)^{K-1} \right] \times \sigma Q \left(\frac{\sqrt{\Delta N \tau_p} (G^* - \beta_0 (\Delta N \tau_p)^{-2/3} - g/\Delta N)}{\sqrt{\alpha_0^2 + g/\Delta N}} \right). \quad (23)$$

We define

$$f(\Delta N) = \frac{\sqrt{\Delta N \tau_p} (G^* - \beta_0 (\Delta N \tau_p)^{-2/3} - g/\Delta N)}{\sqrt{\alpha_0^2 + g/\Delta N}}.$$

The derivative of $f(\Delta N)$ with respect to τ_p can be expressed as follows:

$$\begin{aligned} \frac{\partial f(\Delta N)}{\partial \Delta N} &= \frac{\frac{1}{2} \Delta N^{-\frac{1}{2}} \tau_p^{\frac{1}{2}} (G^* - g/\Delta N) + \frac{1}{6} \beta_0 \Delta N^{-\frac{7}{6}} \tau_p^{-\frac{1}{6}}}{2(\alpha_0^2 + g/\Delta N)^{3/2}} \\ &\geq 0 \end{aligned}$$

which means that P_M is monotonically nonincreasing function of ΔN .

Therefore, we can formulate an optimization problem to investigate how to allocate the pilots and access latency can meet diverse AFP requirements. The main design objective is to find the minimum access latency in terms of time slots, while guaranteeing that the UEs are successfully resolved with the AFP below the required threshold P_M^* . Thus, the optimization problem can be formulated as follows:

$$\min_{\Delta N^*, \tau_p^*} \Delta N^* \quad (24a)$$

$$\text{s.t. } P_M(\tau_p^*, \Delta N^*) \leq P_M^* \quad (24b)$$

$$0 < \tau_p^* \leq \tau_p^- \quad (24c)$$

$$\Delta N^* \in \mathbb{R}. \quad (24d)$$

The objective function (24a) denotes the minimum access latency for the transmission in a frame, (24b) denotes the AFP constraint to ensure that the UEs can access the system reliably, and (24c) denotes the preallocate pilots constraint. For the sake of complete analysis, we assume that ΔN^* is a continuous variable in the field of real numbers. Since the objection function and standardized constraint functions are convex functions over ΔN^* and τ_p^* , the closed-form solution of the above optimization problem can be obtained by using the KKT conditions.

The corresponding Lagrangian function for ΔN^* optimization can be expressed as

$$\begin{aligned} L(\tau_p^*, \Delta N^*, \lambda_1, \lambda_2, \lambda_3) &= \Delta N^* + \lambda_1 (P_M(\tau_p^*, \Delta N^*) - P_M^*) \\ &\quad + \lambda_2 (-\tau_p^*) + \lambda_3 (\tau_p^* - \tau_p^-). \end{aligned} \quad (25)$$

Then, we can write the KKT conditions as follows:

$$\begin{cases} \lambda_i^* \geq 0, i = 1, 2, 3 \\ \lambda_1^* (P_M(\tau_p^*, \Delta N^*) - P_M^*) = 0 \\ \lambda_2^* (-\tau_p^*) = 0 \\ \lambda_3^* (\tau_p^* - \tau_p^-) = 0 \\ \frac{\partial L}{\partial \tau_p^*} = \lambda_1^* \cdot \frac{\partial P_M(\tau_p^*, \Delta N^*)}{\partial \tau_p^*} - \lambda_2^* + \lambda_3^* = 0 \\ \frac{\partial L}{\partial \Delta N^*} = 1 + \lambda_1^* \cdot \frac{\partial P_M(\tau_p^*, \Delta N^*)}{\partial \Delta N^*} = 0. \end{cases} \quad (26)$$

Since $([\partial P_M(\tau_p^*, \Delta N^*)]/[\partial \Delta N^*]) < 0$ and $(\partial L)/(\partial \Delta N^*) = 1 + \lambda_1^* \cdot ([\partial P_M(\tau_p^*, \Delta N^*)]/[\partial \Delta N^*]) = 0$, hence $\lambda_1^* > 0$. Also, we can show that $(P_M(\tau_p^*, \Delta N^*) - P_M^*) = 0$ and $\lambda_2^* = 0$ due to $\lambda_1^* > 0$ and $\tau_p > 0$. As proved in Appendix B, $P_M(\tau_p, \Delta N)$ decreases rapidly with the increase of τ_p for $\tau_p \in (0, \min(\tau_p^-, g\tau_p^-))$. Thus, $([\partial P_M(\tau_p^*, \Delta N^*)]/[\partial \tau_p^*]) < 0$ for $\tau_p \in (0, \min(\tau_p^-, g\tau_p^-))$, and $\lambda_3^* = -\lambda_1^* \cdot ([\partial P_M(\tau_p^*, \Delta N^*)]/[\partial \tau_p^*]) > 0$ follows. With some algebraic manipulations, the optimal solutions for ΔN^* and τ_p^* can be obtained as follows:

$$\tau_p^* = \lceil \min(\tau_p^-, g\tau_p^-) \rceil \quad (27)$$

and

$$\Delta N^* = \lceil \left\{ \Delta N^* \mid P_M(\tau_p^*, \Delta N^*) = P_M^* \right\} \rceil \quad (28)$$

where $\lceil \cdot \rceil$ denotes the ceiling function.

It is worth noting that in the low system load regime (i.e., $g < 1$), the optimal value of τ_p in the MSPA protocol is related with system load g to guarantee the UEs to achieve the desired P_M^* . When the system load g is larger than 1, the AFP dramatically increases with g , thus, we should increase ΔN to ensure the desired AFP threshold with the limit of available pilots.

B. Maximization of the System Throughput

From the above analyses, we can find out that the AFP P_M would dramatically increase with the increasing of the system load g when the number of UEs is larger than that of pilots (i.e., $g > 1$). Therefore, to maintain the desired AFP P_M^* under the optimal number of pilots τ_p^* in (27), the access latency ΔN should increase as shown in (23). Considering that the system throughput T is related to ΔN , by substituting (23) into (22), the system throughput can be expressed as follows:

$$\begin{aligned} T &= \frac{K}{\Delta N} \left\{ 1 - \left[1 - \left(1 - \frac{1}{\tau_p} \right)^{K-1} \right] \right. \\ &\quad \left. \times \sigma Q \left(\frac{\sqrt{\Delta N \tau_p} (G^* - \beta_0 (\Delta N \tau_p)^{-2/3} - g/\Delta N)}{\sqrt{\alpha_0^2 + g/\Delta N}} \right) \right\}. \end{aligned} \quad (29)$$

Then, we can maximize the system throughput under P_M^* and τ_p^* by increasing ΔN . Let T^* denote the system throughput with the minimum access latency ΔN^* in (28). We define T_c as the system throughput when the access latency increases from

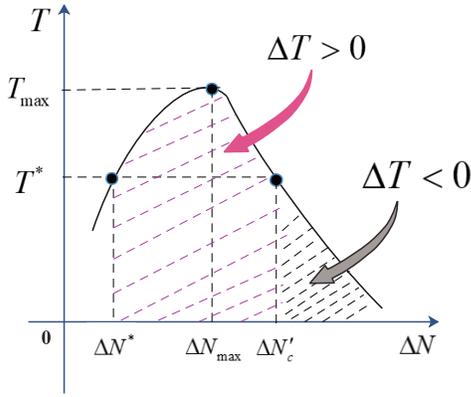


Fig. 7. Variation of the system throughput with the access latency.

ΔN^* to ΔN_c , where $0 < \Delta N^* < \Delta N_c$. The gap between T^* and T_c is denoted by ΔT , which can be evaluated as follows:

$$\begin{aligned} \Delta T &= \frac{K(1 - P_M(\Delta N_c))}{\Delta N_c} - \frac{K(1 - P_M(\Delta N^*))}{\Delta N^*} \\ &= \left[\Delta N^* - \Delta N_c + \Delta N_c P_M(\Delta N^*) \right. \\ &\quad \left. - \Delta N^* P_M(\Delta N_c) \right] \frac{K}{\Delta N_c \Delta N^*} \end{aligned} \quad (30)$$

where $P_M(\Delta N^*) = P_M(\Delta N^*, \tau_p^*)$ and $P_M(\Delta N_c) = P_M(\Delta N_c, \tau_p^*)$, for the fixed τ_p^* .

Define $\Delta f = \Delta N^* - \Delta N_c + \Delta N_c P_M(\Delta N^*) - \Delta N^* P_M(\Delta N_c)$ as the first item in (30). Considering that $\Delta N_c \Delta N^* > 0$ and $K > 0$, then $\Delta T > 0$ if and only if $\Delta f > 0$. Therefore, two cases can be distinguished.

Case 1 $\Delta f < 0$: The condition of $\Delta f < 0$ can be expressed as follows:

$$\Delta N_c > \Delta N^*(1 - P_M(\Delta N_c))/(1 - P_M(\Delta N^*)). \quad (31)$$

Define $\Delta N'_c = \Delta N^*(1 - P_M(\Delta N_c))/(1 - P_M(\Delta N^*))$, where $P_M(\Delta N'_c) = P_M(\Delta N'_c, \tau_p^*)$. From (31), we can find out that the system throughput would deteriorate than T^* if $\Delta N_c > \Delta N'_c$, as the area of $\Delta T < 0$ shown in Fig. 7. When $\Delta N_c = \Delta N'_c$, the system throughput T_c is approximately equal to T^* with the lowest AFP.

Case 2 $\Delta f > 0$: Similarly, the condition of $\Delta f > 0$ can be derived as

$$\Delta N_c \leq \Delta N'_c. \quad (32)$$

Note that the increasing of the access latency from ΔN^* to $\Delta N'_c$ can enhance the system throughput, which is corresponding to the area of $\Delta T > 0$ as shown in Fig. 7. Then, we can find the maximum system throughput T_{\max} for our MSPA random-access protocol. The derivative of ΔT can be expressed as follows:

$$\frac{d\Delta T}{d\Delta N_c} = \frac{-K \left(\Delta N_c \frac{dP_M(\Delta N_c)}{d\Delta N_c} + 1 - P_M(\Delta N_c) \right)}{\Delta N_c^2}. \quad (33)$$

Considering that P_M is monotonically nonincreasing, then T_{\max} can be achieved at $([d\Delta T]/[d\Delta N_c])|_{\Delta N_c = \Delta N_{\max}} = 0$.

Note that $\Delta N^* < \Delta N_{\max} < \Delta N'_c$ as shown in Fig. 6. The numerical solution can be easy to obtain due to ΔN is an integer in practical. It is worth noting that the system throughput

TABLE II
SIMULATION ENVIRONMENT

Parameter	Setting
System load	$0.2 \leq g \leq 1.4$
Radius of each cell	250 m
Distance between the BS and UE	$D \geq 25$ m
Number of available pilots	$20 \leq \tau_p \leq 120$
Number of UEs	$10 \leq K \leq 250$
Requirements of AFP	$P_M^* = [10^{-1}, 10^{-2}, 10^{-3}]$

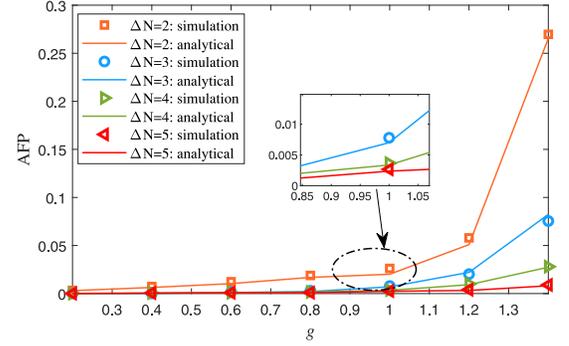


Fig. 8. AFP of the MSPA protocol derived by (14) versus the system load g for different access latency ΔN with $\tau_p = 60$.

improves when the access latency is increasing from ΔN^* to $\Delta N'_c$, and the AFP is also decreasing.

V. SIMULATION RESULTS AND DISCUSSION

In this section, we present the numerical results based on the Monte Carlo simulations and evaluate the theoretical analysis in Section IV. Our MSPA random-access protocol can achieve higher performance in comparison to other relevant pilot allocation random-access protocols, such as SUCR-IPA [14] and SUCR-GBPA [13], in terms of AFP, throughput, and access latency. Moreover, the verification of the optimal transmission parameters derived in Section IV is also presented in the following. For the sake of practical consideration, we consider a crowded random-access scenario in [31], where the radius of each cell is 250 m, a finite number of UEs distribute uniformly in each cell at locations further than 25 m from the BS. We set $10 \leq K \leq 250$ and $20 \leq \tau_p \leq 120$, and a reasonable system load range $0.2 \leq g \leq 1.4$ can be investigated for evaluating the accuracy of the derived AFP. The rest of the simulation parameters is summarized in Table II.

Fig. 8 illustrates the AFP of the proposed MSPA random-access protocol versus the system load g , where $12 \leq K \leq 84$ and $\tau_p = 60$. The simulation results are averaged over a total of 10000 iterations, and match well with the analysis results obtained by (23). As can be observed, for all the considered access latency in Fig. 8, i.e., $\Delta N = 2, 3, 4, 5$, the AFP increases at a slower pace from $g = 0.2$ to $g = 1.2$, and dramatically increases when $g > 1.2$. At $g = 1.4$, more than 27% of the UEs failed their access attempts when $\Delta N = 2$, while only about 2.5% of the UEs failed to access the system for $\Delta N = 4$. Thus, the increase of ΔN allows a significant improvement on the AFP, which indicates that for the UEs

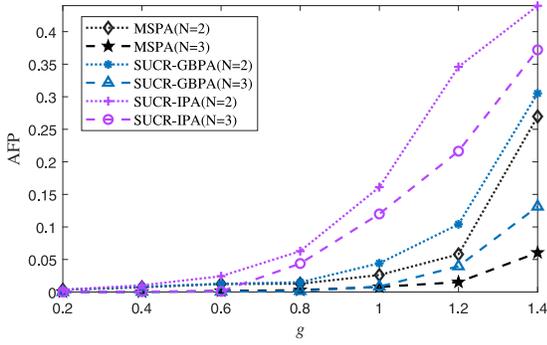


Fig. 9. AFP of MSPA, SUCR-IPA, and SUCR-GBPA protocols versus the system load g for different access latency ΔN with $\tau_p = 40$.

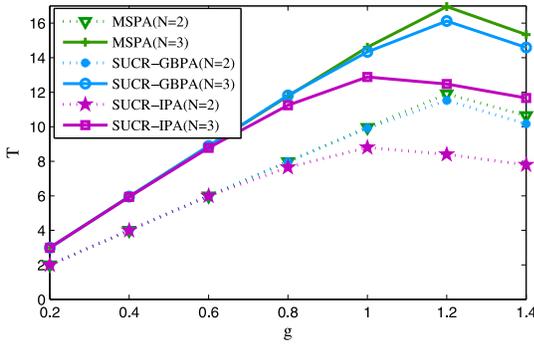
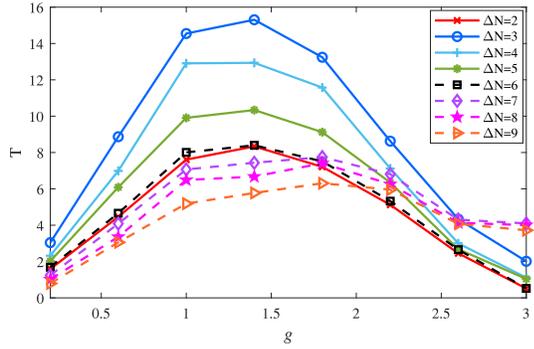


Fig. 10. Throughput of MSPA, SUCR-IPA, and SUCR-GBPA protocols versus the system load g for different access latency ΔN with $\tau_p = 60$. (a) Analytical and simulation results of the throughput for MSPA protocols with diverse value of ΔN . (b) Throughput comparison between the MSPA and relevant RA protocols.

target on low AFP in the extremely crowded random-access scenario, a tradeoff between the access latency and AFP is acceptable.

Fig. 9 shows the AFP comparison of the MSPA, SUCR-IPA, and SUCR-GBPA schemes versus the system load, where $8 \leq K \leq 56$ and $\tau_p = 40$. Note that the MSPA scheme outperforms the SUCR-GBPA and SUCR-IPA schemes under $\Delta N = 2$, i.e., the original setup of access latency in these random-access protocols. In the high system load region, i.e., $g > 1.0$, the failed UEs that have a pilot collision with other UEs in step 1, are permitted to reselect a pilot from the reselection pilot set

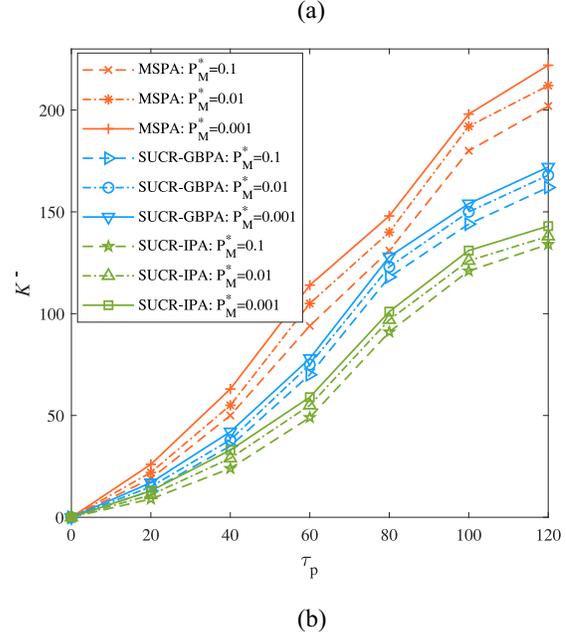
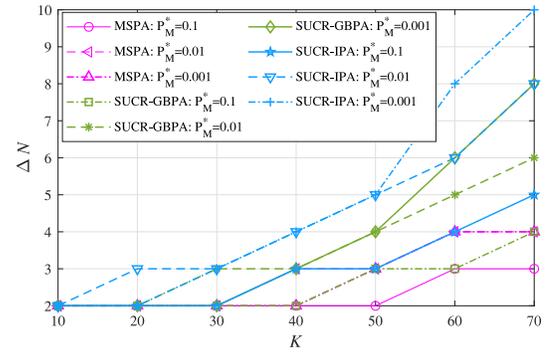


Fig. 11. Optimal value of ΔN^* and the maximum tolerable number of UEs K^* for the MSPA and other random-access schemes under diverse constraints $P_M^* = 10^{-1}$, 10^{-2} , or 10^{-3} . (a) Minimal access latency versus the active UEs K with $\tau_p = 45$. (b) Maximum tolerable number of UEs versus the allocated pilots τ_p with $\Delta N = 3$ and $\tau_p = 0-120$.

\mathbf{R} that consist of the idle pilots and the collision pilots, and send it to the BS in step 3, as shown in Fig. 2.

The simulation results of the system throughput of our MSPA random-access protocol agree well with the theoretical expression in (29) as shown in Fig. 10(a), where $12 \leq K \leq 84$ and $\tau_p = 60$. The system throughput of $\Delta N = 3$ is higher than all the rest values of ΔN due to the fact that 3 is the closest value to approach ΔN_{\max} . Then, the system throughput decreases as $\Delta N > 3$, and approaches to T^* for $\Delta N = \Delta N'_c = 6$, which refers to the case of $\Delta T = 0$ in Section IV-B. Moreover, the peaks of the system throughput with different ΔN are appeared at about $g = 1.4$. As the number of active UEs is an integer in practice, the accurate maximum system throughput can be found by utilizing the linear search methods, e.g., golden section method (GSM) or Fibonacci search method (FSM) [37], [38]. In addition, when $g > 2.5$, the throughput for $\Delta N > 6$ decreases at a slower pace with the increase of system load, while the throughput for $2 \leq \Delta N \leq 6$ decreases dramatically with the increase of system load.

The comparison of system throughput with different ΔN of these random-access protocols is shown in Fig. 10(b). Still, our MSPA scheme outperforms the SUCR-GBPA and SUCR-IPA schemes. Specifically, the SUCR-IPA scheme experiences a saturation at roughly $g = 1.0$, whereas the MSPA and SUCR-GBPA schemes continue to increase until $g \geq 1.2$. The main reason is that the introduced access class barring factor in the SUCR-IPA scheme only permits a part of failed UEs to contend the idle pilots for retransmission, while both the MSPA and the SUCR-GBPA allow all of failed UEs to retransmit, and thus can further benefit the system throughput.

The optimal value of ΔN^* and the tolerable number of UEs K with diverse AFP constraints for our MSPA and other random-access schemes are shown in Fig. 11. In Fig. 11(a), we set $\tau_p^- = 45$ and $P_M^* = [10^{-1}, 10^{-2}, 10^{-3}]$, respectively. For our MSPA random-access protocol, τ_p and ΔN have been optimized by (27) and (28). For the special case that $P_M^* = 10^{-3}$ and $K = 130$, the MSPA protocol can cope with pilot collisions, and keep the AFP below P_M^* by $\Delta N = 4$, which thanks to the pilot reselection step 3. For the mMTC scenario with limited available pilots, we set $\Delta N = 3$ and present the tolerable number of UEs K versus the number of pilots $\tau_p = 20\text{--}120$ for three AFP thresholds $P_M^* = [10^{-1}, 10^{-2}, 10^{-3}]$ in Fig. 11(b). It can be observed that the MSPA scheme can tolerate more UEs in comparison with the SUCR-IPA and SUCR-GBPA schemes when the same values of τ_p and P_M^* are given.

VI. CONCLUSION

In this article, the MSPA random-access protocol has been proposed for crowded mMTC scenarios. We have given the bipartite graph representation of our MSPA protocol, and shown the SIC procedure for resolving pilot collisions. By utilizing the finite-length analysis, we have derived the expressions to the AFP and system throughput for our MSPA random-access protocol, and proposed a guideline to design a satisfying MSPA protocol for the crowded mMTC scenario, i.e., minimization the access latency or maximization the system throughput of the mMTC UEs under diverse AFP constraints. Simulation results show that the investigated MSPA random-access protocol can significantly outperform other relevant random-access protocols in terms of reducing the AFP and improving the system throughput.

APPENDIX A

PROOF OF THE EXPRESSION OF ε

Since the SIC procedure of the MSPA protocol is similar to the iterative erasure decoding for an erasure code over the bipartite graph, there is a mapping between the BP thresholds of the erasure code and that of the MSPA protocol, i.e., $\delta^* = G^*(1 - r)$. The input information m of the erasure code is mapped to the UEs in the MSPA protocol, which is understandable as both the m input information and the K UEs are considered as the unknown message and follow the same BP decoding procedure.

Moreover, the scaling parameters δ_r , α_r , and β_r can be expressed as follows [35]:

$$\varepsilon_r^* = \varepsilon_r^* \frac{1 - r}{1 - r'} \quad (34)$$

$$\alpha_r^* = \alpha_r^* \left(\frac{1 - r}{1 - r'} \right)^{1/2} \quad (35)$$

$$\beta_r^* = \beta_r^* \left(\frac{1 - r}{1 - r'} \right)^{1/3} \quad (36)$$

For the case of $m = n$, that is, the nominal rate $r = 0$, the above expressions can be rewritten as follows:

$$\varepsilon_r^* = \varepsilon_0^*(1 - r) \quad (37)$$

$$\alpha_r^* = \alpha_0^*(1 - r)^{1/2} \quad (38)$$

$$\beta_r^* = \beta_0^*(1 - r)^{1/3} \quad (39)$$

By substituting $\delta^* = G^*(1 - r)$ and $g = \delta\Delta N/(1 - r)$ into (17), we have

$$\varepsilon \approx Q \left(\frac{\sqrt{\Delta N \tau_p} (G^* - \beta_0 (\Delta N \tau_p)^{-2/3} - g/\Delta N)}{\sqrt{\alpha_0^2 + g/\Delta N}} \right). \quad (40)$$

APPENDIX B

PROOF OF PROPOSITION 1

As discussed in Section III, we have

$$P_M = \left[1 - \left(1 - \frac{1}{\tau_p} \right)^{K-1} \right] \times \sigma Q \left(\frac{\sqrt{\Delta N \tau_p} (G^* - \beta_0 (\Delta N \tau_p)^{-2/3} - g/\Delta N)}{\sqrt{\alpha_0^2 + g/\Delta N}} \right).$$

Now, we define $f_1(\tau_p) = [1 - (1 - 1/\tau_p)^{K-1}]$ and

$$f_2(\tau_p) = \sigma Q \left(\frac{\sqrt{\Delta N \tau_p} (G^* - \beta_0 (\Delta N \tau_p)^{-2/3} - g/\Delta N)}{\sqrt{\alpha_0^2 + g/\Delta N}} \right).$$

We observe from the derivative of $f_1(\tau_p)$ and $f_2(\tau_p)$ that $f_1(\tau_p)$ is a monotonically decreasing function for $\tau_p \in (0, \min(\tau_p^-, g\tau_p^-))$ and is a monotonically increasing function for $\tau_p \in (\min(\tau_p^-, g\tau_p^-), \tau_p^-)$. Similarly, $f_2(\tau_p)$ is a monotonically decreasing function for $\tau_p \in (0, \tau_p^-)$. Therefore, P_M decreases rapidly with the increase of τ_p for $\tau_p \in (0, \min(\tau_p^-, g\tau_p^-))$.

REFERENCES

- [1] Y. Wu, X. Gao, S. Zhou, W. Yang, Y. Polyanskiy, and G. Caire, "Massive access for future wireless communication systems," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 148–156, Aug. 2020.
- [2] Z. Lin, M. Lin, T. de Cola, J.-B. Wang, W.-P. Zhu, and J. Cheng, "Supporting IoT with rate-splitting multiple access in satellite and aerial integrated networks," *IEEE Internet Things J.*, early access, Jan. 14, 2021, doi: [10.1109/IJOT.2021.3051603](https://doi.org/10.1109/IJOT.2021.3051603).
- [3] J. Jiao, Y. Sun, S. Wu, Y. Wang, and Q. Zhang, "Network utility maximization resource allocation for NOMA in satellite-based Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 1391–1404, Apr. 2020.
- [4] C. Bockelmann *et al.*, "Towards massive connectivity support for scalable mMTC communications in 5G networks," *IEEE Access*, vol. 6, pp. 28969–28992, 2018.

- [5] R. Wang *et al.*, "Modeling disruption tolerance mechanisms for a heterogeneous 5G network," *IEEE Access*, vol. 6, pp. 25836–25848, 2018.
- [6] X. Liu and X. Zhang, "NOMA-based resource allocation for cluster-based cognitive industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 16, no. 8, pp. 5379–5388, Aug. 2020.
- [7] R. Abbas, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "A novel analytical framework for massive grant-free NOMA," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 2436–2449, Mar. 2019.
- [8] J. Jiao, J. Zhou, S. Wu, and Q. Zhang, "Superimposed pilot code-domain NOMA scheme for satellite-based Internet of Things," *IEEE Syst. J.*, early access, Nov. 6, 2020, doi: [10.1109/JSYST.2020.3032749](https://doi.org/10.1109/JSYST.2020.3032749).
- [9] L. Sanguinetti, A. A. D. Amico, M. Morelli, and M. Debbah, "Random access in massive MIMO by exploiting timing offsets and excess antennas," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6081–6095, Dec. 2018.
- [10] J. H. Sørensen, E. De Carvalho, Č. Stefanović, and P. Popovski, "Coded pilot random access for massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8035–8046, Dec. 2018.
- [11] J. H. Sørensen, E. De Carvalho, and P. Popovski, "Massive MIMO for crowd scenarios: A solution based on random access," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2014, pp. 352–357.
- [12] E. Björnson, E. De Carvalho, E. G. Larsson, and P. Popovski, "Random access protocol for massive MIMO: Strongest-user collision resolution (SUCR)," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2016, pp. 1–6.
- [13] H. Han, Y. Li, and X. Guo, "A graph-based random access protocol for crowded massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7348–7361, Nov. 2017.
- [14] H. Han, X. Guo, and Y. Li, "A high throughput pilot allocation for M2M communication in crowded massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9572–9576, Oct. 2017.
- [15] J. Jiao, L. Xu, S. Wu, Y. Wang, R. Lu, and Q. Zhang, "Unequal access latency random access protocol for massive machine-type communications," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 5924–5937, Sep. 2020.
- [16] H. Jiang, D. Qu, J. Ding, and T. Jiang, "Multiple preambles for high success rate of grant-free random access with massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 18, no. 10, pp. 4779–4789, Oct. 2019.
- [17] S. Gong, *et al.*, "Toward optimized network capacity in emerging integrated terrestrial-satellite networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 263–275, Feb. 2020.
- [18] R. Abbas, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "Random access for M2M communications with QoS guarantees," *IEEE Trans. Commun.*, vol. 65, no. 7, pp. 2889–2903, Jul. 2017.
- [19] B.-Z. Shen and K. B. Cameron, "Stopping and/or reducing oscillations in low density parity check (LDPC) decoding," U.S. Patent 7296216, Nov. 13, 2007.
- [20] D. Sejdinovic, R. J. Piechocki, and A. Doufexi, "AND-OR tree analysis of distributed LT codes," in *Proc. IEEE Inf. Theory Workshop Netw. Inf. Theory*, 2009, pp. 261–265.
- [21] M. Ivanov, F. Brännström, A. G. I. Amat, and P. Popovski, "Error floor analysis of coded slotted ALOHA over packet erasure channels," *IEEE Commun. Lett.*, vol. 19, no. 3, pp. 419–422, Mar. 2015.
- [22] A. G. I. Amat and G. Liva, "Finite-length analysis of irregular repetition slotted ALOHA in the waterfall region," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 886–889, May 2018.
- [23] F. Lazaro and Č. Stefanović, "Finite-length analysis of frameless ALOHA with multi-user detection," *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 769–772, Apr. 2017.
- [24] E. Sandgren, A. G. I. Amat, and F. Brännström, "On frame asynchronous coded slotted ALOHA: Asymptotic, finite length, and delay analysis," *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 691–704, Feb. 2016.
- [25] A. Amraoui, A. Montanari, and R. Urbanke, "Analytic determination of scaling parameters," in *Proc. IEEE Int. Symp. Inf. Theory*, 2006, pp. 562–566.
- [26] E. Björnson and E. Jorswieck, "Optimal resource allocation in coordinated multi-cell systems," *Found. Trends Commun. Inf. Theory*, vol. 9, nos. 2–3, pp. 113–381, 2013.
- [27] J. C. Marinello, T. Abrão, R. D. Souza, E. De Carvalho, and P. Popovski, "Achieving fair random access performance in massive MIMO crowded machine-type networks," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 503–507, Apr. 2020.
- [28] A.-S. Bana *et al.*, "Massive MIMO for Internet of Things (IoT) connectivity," 2019. [Online]. Available: [arXiv:1905.06205](https://arxiv.org/abs/1905.06205).
- [29] H. Masoumi and M. J. Emadi, "Joint pilot and data power control in cell-free massive MIMO system," in *Proc. 5th Int. Conf. Millimeter Wave Terahertz Technol. (MMWaTT)*, 2018, pp. 34–37.
- [30] R. Abbas, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "A multi-layer grant-free NOMA scheme for short packet transmissions," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2018, pp. 1–6.
- [31] E. Björnson, E. De Carvalho, J. H. Sørensen, E. G. Larsson, and P. Popovski, "A random access protocol for pilot allocation in crowded massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2220–2234, Apr. 2017.
- [32] E. Björnson, L. Sanguinetti, and M. Debbah, "Massive MIMO with imperfect channel covariance information," in *Proc. 50th Asilomar Conf. Signals Syst. Comput.*, 2016, pp. 974–978.
- [33] X. Liu, X. B. Zhai, W. Lu, and C. Wu, "QoS-guarantee resource allocation for multibeam satellite industrial Internet of Things with NOMA," *IEEE Trans. Ind. Informat.*, vol. 17, no. 3, pp. 2052–2061, Mar. 2021.
- [34] A. Abdi, W. C. Lau, M.-S. 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.
- [35] A. Amraoui, A. Montanari, T. Richardson, and R. Urbanke, "Finite-length scaling for iteratively decoded LDPC ensembles," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 473–498, Feb. 2009.
- [36] G. Liva, "Graph-based analysis and optimization of contention resolution diversity slotted aloha," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 477–487, Feb. 2011.
- [37] M. Holst, R. E. Kozack, F. Saied, and S. Subramaniam, "Treatment of electrostatic effects in proteins: Multigrid-based Newton iterative method for solution of the full nonlinear Poisson–Boltzmann equation," *Proteins Struct. Function Bioinform.*, vol. 18, no. 3, pp. 231–245, Mar. 1994.
- [38] N. Lopez, M. Nunez, I. Rodriguez, and F. Rubio, "Introducing the golden section to computer science," in *Proc. 1st IEEE Int. Conf. Cogn. Informat.*, 2002, pp. 203–212.



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

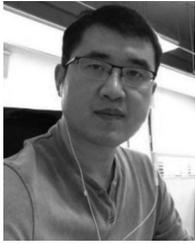
From 2011 to 2015, he was a Postdoctoral Research Fellow with the Communication Engineering Research Centre, Shenzhen Graduate School, HIT, Shenzhen, China. From 2016 to 2017, he was a China Scholarship Council Visiting Scholar with the School of Electrical and Information Engineering, University of Sydney,

Sydney, NSW, Australia. From 2017 to 2019, he was an Assistant Professor with the HIT (Shenzhen), Shenzhen, where he has been an Associate Professor with the Department of Electrical and Information Engineering since 2020. He is also an Associate Professor with Peng Cheng Laboratory, Shenzhen. His current interests include error control codes, satellite communications, and massive random access.



Liang Xu received the M.S. degree in communication engineering from Harbin Institute of Technology (Shenzhen), Shenzhen, China, in 2021.

His research interests include massive MIMO, random access, and satellite communications.



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

From 2009 to 2011, he was a Postdoctoral Fellow with the Department of Electronics and Information Engineering, Shenzhen Graduate School, Harbin Institute of Technology (Shenzhen), Shenzhen, China, where he has been an Associate Professor since 2012. He is also an Associate Professor with Peng Cheng Laboratory, Shenzhen. From 2014 to 2015, he was a Visiting Researcher with BBCR, University of Waterloo, Waterloo, ON, Canada. He holds over 40 Chinese patents. His current research interests include wireless image/video transmission, space communications, advanced channel coding techniques, and B5G wireless transmission technologies. He has authored or coauthored over 100 papers in the above areas.



Rongxing Lu (Fellow, IEEE) received the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, in 2012.

He worked as an Assistant Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, from April 2013 to August 2016. He worked as a Postdoctoral Fellow with the University of Waterloo from May 2012 to April 2013. He is currently an Associate Professor with the Faculty of Computer Science, University of New Brunswick, Fredericton, NB, Canada. His research interests include applied cryptography, privacy-enhancing technologies, and IoT-big data security and privacy. He has published extensively in his areas of expertise.

Dr. Lu was a recipient of the eight best (student) paper awards from some reputable journals and conferences. He was awarded the most prestigious Governor General's Gold Medal from the University of Waterloo; and won the 8th IEEE Communications Society (ComSoc) Asia-Pacific Outstanding Young Researcher Award in 2013. He is the Winner of 2016–2017 Excellence in Teaching Award, FCS, UNB. He is currently a Senior Member of the IEEE Communications Society. He currently serves as the Vice-Chair (Conferences) of IEEE ComSoc Communications and Information Security Technical Committee.



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

From 1999 to 2003, he was an Assistant Professor with the University of Tokushima. From 2003 to 2005, he was an Associate Professor with Shenzhen Graduate School, HIT, Shenzhen, China, and was the Founding Director of the Communication Engineering Research Center, School of Electronic and Information Engineering (EIE). Since 2005, he has been a Full Professor and the Dean of EIE School, HIT. His research interests include aerospace communications and networks, wireless communications and networks, cognitive radios, signal processing, and biomedical engineering.

Dr. Zhang received the National Science Fund for Distinguished Young Scholars, the Young and Middle-Aged Leading Scientist of China, the Chinese New Century Excellent Talents in University, and the Three Scientific and Technological Awards from Governments. He is on the Editorial Board of some academic journals, such as the *Journal of Communication*, the *KSII Transactions on Internet and Information Systems*, and *Science China Information Sciences*. He has been a TPC Member of INFOCOM, ICC, GLOBECOM, and WCNC, and other flagship conferences in communications. He was the Associate Chair for Finance of the ICMMT 2012, the TPC Co-Chair of the IEEE/CIC ICC 2015, and the Symposium Co-Chair of the CHINACOM 2011 and the IEEE VTC 2016 (Spring). He was the Founding Chair of the IEEE Communications Society Shenzhen Chapter.