# Performance Analysis of Resource Hopping-Based Grant-Free Multiple Access for Massive IoT Networks

Young-Seok Lee, *Student Member, IEEE*, Ki-Hun Lee, *Student Member, IEEE*,
Han Seung Jang, *Member, IEEE*, Gweondo Jo, and Bang Chul Jung, *Senior Member, IEEE*

*Abstract*—In this letter, we mathematically analyze the bit-error-rate (BER) and frame-error-rate (FER) performance of the resource hopping-based grant-free multiple access (RH-GFMA) systems that were recently proposed for 6G massive Internet-of-Things (mIoT) networks. In the RH-GFMA systems, each IoT device intermittently sends a short-packet to the access point (AP) based on a predetermined resource hopping pattern (HP). Each IoT device may experience HP *collisions* within the packet, but the effect of HP collisions on error performance can be reduced to be *collision-friendly* by exploiting multi-antenna-based signal detection and error-correction capability of channel codes. In this letter, hence, we rigorously analyze the effect of HP collisions on the error performance in the RH-GFMA system where repetition codes (RC) and tail-biting convolutional codes (TB-CC) are considered channel coding schemes for satisfying low-complexity and low-latency requirements of mIoT networks. Through computer simulations, we validate that our mathematical analysis is well-matched with the simulation results.

*Index Terms*—6G, grant-free multiple access (GFMA), massive machine-type communication (mMTC), massive Internet-of-Things (mIoT), resource hopping (RH), bit-error-rate (BER).

## I. INTRODUCTION

THE SIXTH-GENERATION (6G) communications are expected to become one of the core infrastructures to support vertical industries such as building and factory automation, manufacturing, etc [1]. Massive Internet-of-Things (mIoT)-based applications are attracting significant attention since they support seamless automatic management and connectivity without human intervention [2]. Such 6G IoT applications require higher reliability and lower latency than 5G massive machine-type communication (mMTC) applications. They are called *critical* mMTC since they require

not only massive connectivity but ultra-high reliability and low-latency [3], [4].

A *two-step* random access (RA) technique has been applied to the 5G new radio (NR) standard (Release 16) to reduce signaling overhead and latency by abandoning scheduling request and grant procedures in the conventional RA procedure [5], which can be regarded as grant-free multiple access (GFMA) technique in the sense that an uplink device sends packet without a grant from an access point (AP). Hence, grant-free non-orthogonal multiple access (GF-NOMA) techniques have been actively investigated to resolve the effect of collisions in GFMA systems [6], [7]. However, in future 6G mIoT networks such as critical mMTC scenarios in which massive IoT devices are accommodated with the same quality-of-service (QoS) level, it is difficult to completely suppress resource collisions between IoT devices. Therefore, even if collisions are allowed to some extent, massive access techniques that sufficiently satisfy the required QoS are required. In other words, *collision-friendly* GF-NOMA techniques are required [4].

In this regard, a novel resource hopping-based GFMA (RH-GFMA) framework was recently proposed to sufficiently alleviate the effect of resource collisions and to guarantee a delay constraint in mIoT networks [8]. The low-density parity-check (LDPC) code was used as a channel coding scheme to improve reliability [8]. However, modern channel codes such as LDPC, polar, and turbo codes may not be feasible for practical mIoT networks.[1] In addition, the error performance of the RH-GFMA was not mathematically characterized even though the effect of resource collisions on bit-error-rate (BER) or frame-error-rate (FER) is a fundamental performance indicator in RH-GFMA systems. Therefore, our main contributions in this letter are summarized as follows

- We mathematically analyze the BER and FER performance of the channel-coded RH-GFMA system considering repetition code (RC) and tail-biting convolutional code (TB-CC), respectively. To the best of our knowledge, the coded BER analysis for the RH-GFMA system considering resource collisions has not been investigated in the existing literature.
- We also analyze the asymptotic BER performance for the high signal-to-noise ratio (SNR) regime to intuitively characterize system parameters affecting the BER performance of the RH-GFMA system. These results

Young-Seok Lee, Ki-Hun Lee, and Bang Chul Jung are with the Department of Electronics Engineering, Chungnam National University, Daejeon 34134, South Korea (e-mail: yslee@o.cnu.ac.kr; kihun.h.lee@cnu.ac.kr; bcjung@cnu.ac.kr).

Han Seung Jang is with the School of Electrical, Electronic Communication, and Computer Engineering, Chonnam National University, Yeosu 59626, South Korea (e-mail: hsjang@chonnam.ac.kr).

Gweondo Jo is with the Future Mobile Communication Research Division, Electronics and Telecommunications Research Institute, Daejeon 34129, South Korea (e-mail: gdjo@etri.re.kr).

[1]A simple but effective channel code is required for practical mIoT networks since each IoT device sporadically sends short packets and the AP needs to operate in a low-complexity and low-latency manner [10]. Thus, LDPC or turbo code is not appropriate since they adopt an iterative decoding architecture that yields large hardware and computation complexity, and a large decoding delay at the receiver. Moreover, it has been known that they have significant performance gaps when considering short block lengths on the order of a few hundred bits [11].
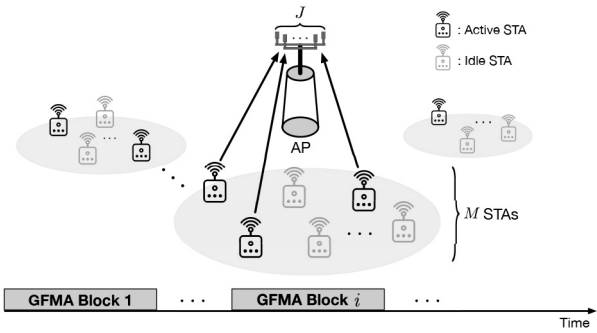
Fig. 1. System model of the RH-GFMA system including several GFMA blocks accommodating massive STAs and an AP with $J$ antennas.



Fig. 2. An example of Latin square group-based resource hopping and resource collisions when $L = 2$ and $R = 5$.

allow us to figure out the fundamental characteristics of the RH-GFMA system.

- We validate that our mathematical analysis is well-matched with computer simulations. It is expected that the analytical results can be exploited for network optimization problems to maximize the number of acceptable STAs in the mIoT networks, by elaborately adjusting the system parameters.

## II. SYSTEM MODEL OF RH-GFMA

We consider a single-cell mIoT network where IoT devices (stations: STAs) are equipped with a single antenna and an access point (AP) is equipped with $J$ multiple antennas. As illustrated in Fig. 1, we define a GFMA block consisting of frequency-time resource blocks (RBs) to efficiently accommodate a massive number of STAs [8]. Each GFMA block is assigned to pre-scheduled $M$ STAs, and each STA sends a packet to the AP intermittently by using its own *hopping pattern* (HP) through the GFMA block. The AP is assumed to detect the signal from multiple STAs with a well-known linear zero-forcing (ZF) method for low-complexity and low-latency at the receiver.

Let $\nu$ ($\in [0, 1]$) be the activation probability of each STA, which indicates the probability that an STA sends a packet for a given GFMA block. It is assumed that each STA is independently activated with the same probability $\nu$ in a GFMA block. We define a GFMA block as $R \times R$ time-frequency RBs and a loading factor as $L$ as in [8]. Then, the maximum number of STAs for each GFMA block is given by $M = R \cdot L$, where the loading factor denotes the ratio of the number of assigned STAs for a given GFMA block to the number of available RBs. For example, when $L = 1$, the number of assigned STAs to a GFMA block is equal to the number of available resources, i.e., $R = M$, resource collisions do not occur since each resource can be assigned to each STA exclusively. On the other hand, when $L = 2$, the number of assigned STAs to a GFMA block is twice the number of available resources, i.e., $M = 2R$, resource collisions between two STAs may occur.

We assume that a *unique* preamble index is assigned to each STA based on its IoT device identifier (device-ID) and thus there exist no preamble collisions among STAs. In other words, the AP is assumed to know which STAs send their packets through preamble detection. It is worth noting that each preamble index is mapped to a unique resource HP in the RH-GFMA system [8]. An active STA sends both the preamble and data packet in the *grant-free* manner. In a GFMA block, the data packet is sent over the time-frequency resources determined by the assigned resource HP. After receiving the preambles, the AP first detects the activated STAs and their channel state information (CSI). Then, the AP decodes data
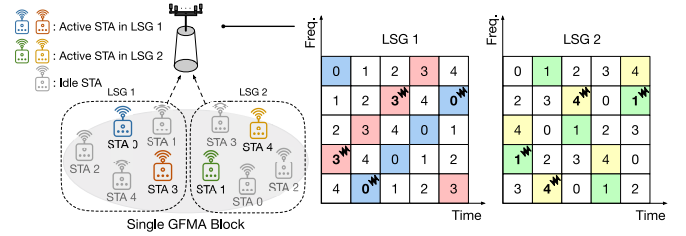
packets of the activated STAs based on the mapping rule between the preamble indices and HPs. We assume that the CSIs of active STAs are exactly known to the AP.

### A. Resource Hopping Mechanism

We adopt the Latin square group (LSG)-based resource hopping method since it results in the best performance among several resource hopping schemes [8]. The set of Latin square matrices is defined as $R \times R$ square matrices $Q^l$ ($l = 1, \ldots, L$), and the element in the $i$-th row and $j$-th column of the $l$-th Latin square matrix is defined as

$$[Q^l]_{i,j} = (l \times i + j) \bmod R, \tag{1}$$

where $[\,\cdot\,]_{i,j}$ denotes the $(i, j)$ entry of the matrix and mod represents the modulo operator. An LSG is defined as a set of STAs that send packets over RBs according to the same Latin square matrix. In the RH-GFMA system, (1) indicates the STA index belonging to the $l$-th LSG and being assigned to the $i$-th frequency and $j$-th time resource. In practice, a single RB may consist of multiple subcarriers and OFDM symbols as in 3GPP LTE and 5G NR systems.

Fig. 2 shows an example of the set of Latin square matrices when $L = 2$ and $R = 5$. There is no collision among STAs belonging to the same LSG regardless of their activity since RBs are exclusively assigned to STAs in the same Latin square matrix. However, resource collisions happen if multiple STAs belonging to different LSGs are activated simultaneously. For a certain RB, the maximum number of colliding STAs is limited by the loading factor $L$. Thus, the probability that a certain RB is collided by $K$ ($\leq L-1$) interfering STAs is given by

$$\alpha_K = \binom{L-1}{K} \nu^K (1-\nu)^{L-1-K}, \tag{2}$$

which is called a *collision probability* in this letter. The case when $K = 0$ means the probability that only a single STA is activated among all LSGs.

### B. Received Signal Model With Zero-Forcing Decorrelator

Without loss of generality, we focus on decoding the received signal of the STA 0 in LSG 1 and we assume the STA 0 in LSG 1 is activated. Thus, we only consider a certain RB assigned to STA 0 in LSG 1 according to resource HP. Let $\mathbf{h}_0 \in \mathbb{C}^{J \times 1}$ and $x_0 \in \mathbb{C}$ denote the wireless channel from the desired STA to the AP and the desired modulated symbol for the certain RB, respectively. Likewise, $\mathbf{h}_k \in \mathbb{C}^{J \times 1}$ and $x_k \in \mathbb{C}$ denote the wireless channel from the STA in LSG $k$ which utilizes the same RB with the STA 0 in LSG 1 to the AP and its modulated symbol, respectively. All elements of $\mathbf{h}_k$ ($0 \leq k \leq L - 1$) are assumed to follow independently and identically distributed (i.i.d.) complex Gaussian distribution with a zero mean and unit variance, i.e., $\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_J)$

where $\mathbf{I}_J$ means the $J \times J$ identity matrix.[2] Then, the received signal at the AP is given by

$$\mathbf{y} = \mathbf{h}_0 x_0 + \sum_{k=1}^{L-1} \mathbf{h}_k b_k x_k + \mathbf{w}, \qquad (3)$$

where $b_k$ denotes the indicator variable which is equal to 1 when the STA in LSG $k$ is active, but otherwise it is equal to 0. Since all STAs have the same activation probability $\nu$, $b_k$ is defined as 1 with a probability of $\nu$ or 0 with a probability of $1 - \nu$. The term $\mathbf{w}$ denotes the additive white Gaussian noise (AWGN) at the AP and it is assumed that $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, N_0\mathbf{I}_J)$.

The log-likelihood ratio (LLR) computation with a joint maximum-likelihood (JML) estimator is optimal [12], but its computational complexity increases exponentially according to the number of colliding STAs in the RH-GFMA system [8]. In order to reduce the computational complexity and decoding latency at the receiver, we adopt a well-known linear ZF decorrelator at the receiver. Let $\mathbf{H} \triangleq [b_0\mathbf{h}_0 \ b_1\mathbf{h}_1 \cdots b_{L-1}\mathbf{h}_{L-1}] \in \mathbb{C}^{J \times (K+1)}$, where $K = \sum_{k=1}^{L-1} b_k$. Then, the estimate for the desired signal $x_0$ after the ZF decorrelator is given by $\widetilde{y} = z\mathbf{a}_{\mathsf{ZF}}\mathbf{y} = zx_0 + \widetilde{w}$, where $\mathbf{a}_{\mathsf{ZF}}$ denotes the first row of the pseudo-inverse matrix of $\mathbf{H}$ and $z = \sqrt{1/[(\mathbf{H}^{\mathsf{H}}\mathbf{H})^{-1}]_{1,1}}$. We regularize the decorrelated signal by $z$ so that the effective noise $\widetilde{w}$ has the same variance of $N_0$ regardless of both wireless channel vectors and the number of colliding STAs.

## III. ERROR PERFORMANCE ANALYSIS

In this section, we mathematically analyze the BER performance of the RH-GFMA system with RC and TB-CC as channel coding schemes that have been used in long-term evolution (LTE) narrowband IoT (NB-IoT) standards. We assume that all STAs send data through quadrature phase-shift keying (QPSK) modulation for low complexity and high energy-efficiency of IoT devices [13]. In addition, we follow the QPSK symbol mapping rule in which the first and second bits of the QPSK symbol correspond to imaginary and real parts, respectively. Then, the bits 0 and 1 are mapped to positive and negative axes on the in-phase and quadrature (I-Q) plane, respectively.

### A. BER of RH-GFMA With Repetition Code

We first consider the RC as the simplest channel coding scheme since it shows a fundamental framework to analyze the BER of the RH-GFMA system with general channel coding schemes. With $n$-RC, the $i$-th received signal of STA 0 in LSG 1 at the AP with $J$ receive antennas, is given by

$$\mathbf{y}_i (\in \mathbb{C}^{J \times 1}) = \mathbf{h}_0^i x_0 + \sum_{k=1}^{L-1} \mathbf{h}_k^i b_k^i x_k^i + \mathbf{w}_i, \ (1 \le i \le n), \ (4)$$

where $\mathbf{h}_0^i$, $\mathbf{h}_k^i$, $b_k^i$, $x_k^i$, and $\mathbf{w}_i$ are identically defined as in (3) by considering the $i$-th signal among $n$ repetition signals. The desired signal $x_0$ does not change for varying $i$ since the same signal is repeated $n$ times. The number of colliding STAs for the $i$-th signal of STA 0 in LSG 1 is given by $K_i = \sum_{k=1}^{L-1} b_k^i$ and it independently changes for varying $i$.

Moreover, the set of colliding STAs becomes different as $i$ changes. Let $\mathbf{H}_i \triangleq [b_0^i\mathbf{h}_0^i \ b_1^i\mathbf{h}_1^i \cdots b_{L-1}^i\mathbf{h}_{L-1}^i] \in \mathbb{C}^{J \times (K_i+1)}$. With the ZF decorrelator, the estimate for $x_0$ at the $i$-th signal, $\widetilde{y}_i (1 \le i \le n)$, is given by $\widetilde{y}_i = z_i\mathbf{a}_{\mathsf{ZF}}^i\mathbf{y}_i = z_ix_0 + \widetilde{w}_i$, where $\mathbf{a}_{\mathsf{ZF}}^i$ denotes the first row of the pseudo-inverse matrix of $\mathbf{H}_i$ and $z_i = \sqrt{1/[(\mathbf{H}_i^{\mathsf{H}}\mathbf{H}_i)^{-1}]_{1,1}}$. Collecting $n$ repeated signals, it is written as $\widetilde{\mathbf{y}} = \mathbf{z}x_0 + \widetilde{\mathbf{w}}$, where $\mathbf{z} (= [z_1, \ldots, z_n] \in \mathbb{R}^{1 \times n})$ and $\widetilde{\mathbf{w}} (= [\widetilde{w}_1, \ldots, \widetilde{w}_n] \in \mathbb{C}^{1 \times n})$ denote regularization factor vector and the effective noise vector for $n$ repeated signals, respectively.

With the symmetry of QPSK constellations, we assume that the transmitted signal of the desired STA is 00 and we focus on the first bit without loss of generality. Then, the corresponding QPSK symbol becomes $x_0 = \sqrt{E_s/2} + j\sqrt{E_s/2}$. For given $\mathbf{z}$ with $n$-RC, the conditional BER is given by

$$\Pr(\mathbf{s_c} \to \mathbf{s_e}|\mathbf{z}) = Q\left(\sqrt{\frac{E_s \sum_{i=1}^n z_i^2}{N_0}}\right), \qquad (5)$$

where $\mathbf{s_c} = \sqrt{E_s/2}\mathbf{z}$ and $\mathbf{s_e} = -\sqrt{E_s/2}\mathbf{z}$. Let $X = \sum_{i=1}^n z_i^2$ and the probability density function (PDF) of $X$ is given by

$$f_X(x) = \frac{1}{(nJ - \sum_{i=1}^n K_i - 1)!} x^{nJ - \sum_{i=1}^n K_i - 1} e^{-x}, \quad (6)$$

where the random variable $X$ follows Chi-square $(\chi^2)$ distribution with $2(nJ - \sum_{i=1}^n K_i)$ degrees of freedom (DoF). From (5) and (6), the conditional BER for a given number of colliding STAs in each repeated signal based on [12] is given

$$\Pr(\mathbf{s_c} \to \mathbf{s_e}|K_1, \ldots, K_n)$$
$$= \frac{1}{2}\left[1 - \sum_{k=0}^{nJ - \sum_{i=1}^n K_i - 1} \binom{2k}{k}\sqrt{\frac{1}{1 + 1/R_c\gamma}}\left(\frac{1}{4R_c\gamma + 4}\right)^k\right], \quad (7)$$

where $R_c(= 1/n)$ denotes the code-rate, $\gamma = E_b/N_0$, and $E_b = nE_s/2$.

The number of colliding STAs at the $i$-th repeated signal depends on the same collision probability of (2) and the maximum number of colliding STAs is limited to $L - 1$ in the RH-GFMA system. Thus, $\Pr(K_1, \ldots, K_n) = \prod_{i=1}^n \alpha_{K_i}$. Finally, the closed-form of the BER of the RH-GFMA with $n$-RC, $P_b^{(\mathsf{RC})}$, is given by

$$P_b^{(\mathsf{RC})} = \sum_{K_1, \ldots, K_n} \Pr(\mathbf{s_c} \to \mathbf{s_e}|K_1, \ldots, K_n) \Pr(K_1, \ldots, K_n)$$
$$= \sum_{K_1=0}^{L-1} \cdots \sum_{K_n=0}^{L-1} \prod_{i=1}^n \alpha_{K_i} \Pr(\mathbf{s_c} \to \mathbf{s_e}|K_1, \ldots, K_n), \quad (8)$$

where $\Pr(\mathbf{s_c} \to \mathbf{s_e}|K_1, \ldots, K_n)$ is derived in (7).

### B. BER of RH-GFMA With Convolutional Code

In this section, we consider the RH-GFMA system with TB-CC. Since deriving an exact BER for CC is known to be mathematically intractable, we obtain an upper bound of the BER. Specifically, a BER upper bound of CC with Viterbi decoder is given by $P_b^{(\mathsf{CC})} < \sum_{d=d_{\text{free}}}^\infty A_d \Pr(d)$, where $A_d$ denotes the total number of nonzero information bits on all weight $d$ codewords, $d_{\text{free}}$ denotes the free distance which is the minimum distance bits between codewords, and $\Pr(d)$ means the pairwise error probability (PEP) for two distinct

codewords with a Hamming distance of $d$ as a union bound of each pairwise error event [14]. Here, the terms $A_d$ and $d_{\text{free}}$ are determined from the CC encoder structure [15]. The PEP $\Pr(d)$ is in general derived by assuming that a transmitter sends all-zero codeword, i.e., $\mathbf{c}_0 = [0, 0, \ldots, 0]$, but a receiver decodes it as other codewords $\mathbf{c}_d$ with Hamming weight $d$. Considering the QPSK constellation, the conditional PEP of the RH-GFMA system with CC is given by

$$\Pr(d|\mathbf{z}) = \Pr(\mathbf{c}_0 \to \mathbf{c}_d|\mathbf{z}) = Q\left(\sqrt{2R_c \frac{E_b \sum_{i=1}^{d} z_i^2}{N_0}}\right). \quad (9)$$

Following the same derivation process of (8), the upper bound of the coded BER of the RH-GFMA with CC is obtained as

$$P_b^{(\text{CC})} < \sum_{d=d_{\text{free}}}^{\infty} A_d \Pr(d) = \sum_{d=d_{\text{free}}}^{\infty} A_d \, \mathbb{E}_X\left[Q\left(\sqrt{2R_c \frac{E_b \sum_{i=1}^{d} z_i^2}{N_0}}\right)\right]$$

$$= \frac{1}{2}\left[\sum_{d=d_{\text{free}}}^{\infty} A_d \sum_{K_1=0}^{L-1} \cdots \sum_{K_d=0}^{L-1} \prod_{i=1}^{d} \alpha_{K_i}\right.$$

$$\left. \times \left[1 - \sum_{k=0}^{dJ-\sum_{j=1}^{d} K_j - 1} \binom{2k}{k} \sqrt{\frac{1}{1 + 1/R_c\gamma}}\left(\frac{1}{4R_c\gamma + 4}\right)^k\right]\right]. \quad (10)$$

### C. Asymptotic Behavior of BER Performance

Let's revisit (7) for RC and (9) for TB-CC, respectively. Since bit error events in the high SNR region dominate when all terminals collide for all resource blocks, we can state $K_i = L-1, \forall i$. Hence, using the Taylor series expansion under the high SNR region, an asymptotic BER of the RH-GFMA with RC can be derived from (8) as follows:

$$P_b^{(\text{RC})} \approx \frac{\beta_{n,L,J} \cdot \nu^{n(L-1)}}{2R_c^{n(J-L+1)}} \cdot \gamma^{-n(J-L+1)}, \quad (11)$$

where $\beta_{n,L,J}$ denotes the first coefficient of the Taylor series expansion in (7), which can be numerically obtained according to the code-rate, the loading factor, and the number of receive antennas at the AP. In the same manner, an asymptotic BER of the RH-GFMA with TB-CC can also be derived as

$$P_b^{(\text{CC})} \lesssim \frac{A_{d_{\text{free}}} \cdot \beta_{d_{\text{free}},L,J} \cdot \nu^{d_{\text{free}}(L-1)}}{2R_c^{d_{\text{free}}(J-L+1)}} \cdot \gamma^{-d_{\text{free}}(J-L+1)}, \quad (12)$$

because the PEP of the path with distance $d_{\text{free}}$ dominates the bit error events in the Viterbi decoder under the high SNR regime. These expressions allow us to more intuitively obtain the BER behavior with respect to the code-rate $R_c$, the number of receive antennas $J$, and the loading factor $L$.

### D. FER Analysis of RH-GFMA From the BER Analysis

In communication systems, the FER is another important key performance indicator because the frame even with a single bit error is discarded at the receiver. From the definition of FER, which treats as error events for all bits in a frame even if only one bit is confused within a frame, the FER can be approximated from the BER as

$$\text{FER} \approx 1 - (1 - P_b)^B \approx P_b \cdot B, \quad (13)$$

where $P_b$ denotes the bit-error probability with RC or TB-CC and the $B$ denotes the frame length [16]. Therefore, we can elaborately predict the approximated FER performance by thoroughly verifying the analyzed BER performance.
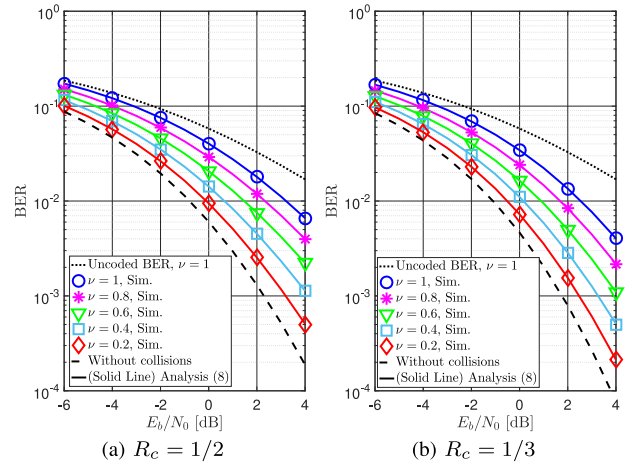


Fig. 3. BER performance of the RH-GFMA system with RC.

### IV. SIMULATION RESULTS

We verify analytical results on the coded BER of the RH-GFMA system through Monte-Carlo simulations. We consider two code-rates for both RC and TB-CC, i.e., $R_c = 1/2$ and $R_c = 1/3$. For the CC encoder, we consider two typical generator polynomials, which are $(133, 171)$ and $(557, 663, 711)$ in octal numbers. Their constraint lengths of the CC encoder are 7 and 9, respectively [15]. We consider a short-packet transmission for critical mMTC scenarios, and the packet is assumed to contain 36 bits. We assume that the QPSK modulation is used. To satisfy the low-latency requirement of delay-critical applications, a concept of *mini-slot* has been proposed in 5G NR, where a single RB can consist of 12 subcarriers and only one OFDM symbol [17]. Assuming that the bandwidth of a single subcarrier is 15kHz, each RB occupies 180kHz in the frequency domain and $72\mu$sec in the time domain [18]. We assume $R = 3$ and $R = 5$ when $R_c = 1/2$ and $R_c = 1/3$, respectively. Then, a single GFMA block spans 540kHz in frequency and 0.216msec in time when $R_c = 1/2$, while it spans 900kHz in frequency and 0.36msec in time when $R_c = 1/3$. The number of antennas at the AP is assumed to be 4 and the loading factor is set to 3. The maximum number of colliding STAs in an RB is limited to 2.

Fig. 3 shows the BER performance of the RH-GFMA system with RCs with (a) $R_c = 1/2$ and (b) $R_c = 1/3$, respectively, for varying $\nu$. When $\nu = 1$, all STAs in the GFMA block send packets through their resource HPs and the RH-GFMA system with RC achieves coding gains of about 2.2 dB when $R_c = 1/2$, and 2.9 dB when $R_c = 1/3$, compared with the uncoded case. It is worth noting that the analysis is matched well with the simulation results.

Figs. 4 and 5 show the BER performance of the RH-GFMA system with TB-CC when $R_c = 1/2$ and $1/3$ for varying $\nu$, respectively. If the latency requirement is set to 1sec, the BER requirement is set to $10^{-4}$, and $R_c = 1/3$, then the RH-GFMA system with TB-CC that occupies bandwidth of 900kHz can accommodate 41,655 STAs regardless of activation probabilities by increasing transmit power at STAs with 3.72dB compared with the case without collisions, respectively. In other words, the analytical results can be exploited for network optimization problems to maximize the number of acceptable STAs in mIoT networks by elaborately adjusting the system parameters, e.g., the loading factor, the number of receive antennas, the required transmission power, etc. In addition, Fig. 5 compares the BER performance of TB-CC with the
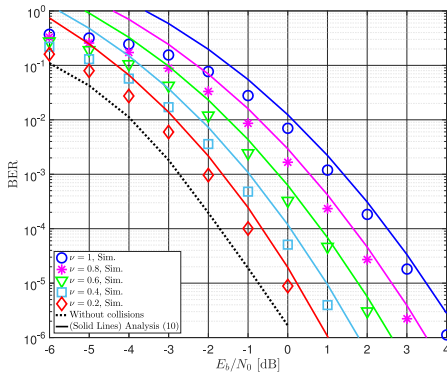
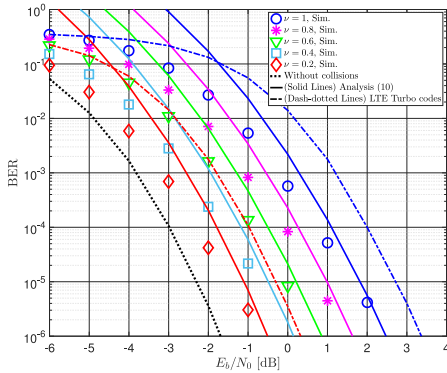Fig. 4. BER performance of the RH-GFMA system with TB-CC when $R_c = 1/2$.



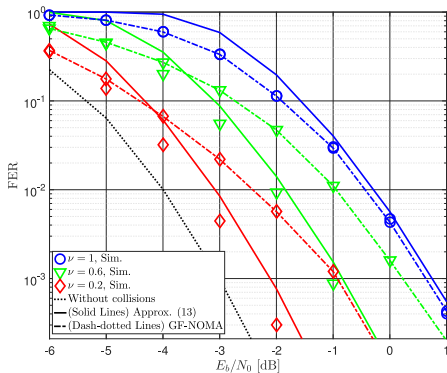Fig. 5. BER performance of the RH-GFMA system with TB-CC when $R_c = 1/3$.



Fig. 6. FER performance of the RH-GFMA system with TB-CC when $R_c = 1/3$.

LTE turbo codes in the RH-GFMA system. It is worth noting that the TB-CC significantly outperforms the LTE turbo codes in terms of BER of the RH-GFMA system especially when considering the short-packet traffic and resource collisions in mIoT networks.

Fig. 6 shows the FER performance of the RH-GFMA system with TB-CC when $R_c = 1/3$ for varying $\nu = 0.2, 0.6$ and 1. This figure verifies that the proposed FER approximation for the RH-GFMA matches well with simulation results. Furthermore, we compare the proposed RH-GFMA with the existing GF-NOMA technique in terms of FER. In the GF-NOMA, STAs in the same GFMA block send their packets through the common resource blocks, where the maximum number of STAs sharing the same resources is limited to $L$

for a fair comparison. When $\nu = 1$, all packets in GF-NOMA and all resources in RH-GFMA collide with all STAs, and thus both systems yield the same FER performance. It is worth noting that the proposed RH-GFMA outperforms the conventional GF-NOMA as the activation probability of STAs decreases since the proposed RH-GFMA efficiently average out the interference due to collisions.

## V. CONCLUSION

In this letter, we mathematically analyzed the channel-coded BER and FER of the RH-GFMA system with linear ZF MIMO detectors at the receiver, which are the first theoretical result for RH-GFMA systems in the literature. We considered repetition and tail-biting convolutional codes as channel coding schemes for satisfying low-complexity and low-latency requirements in massive IoT networks. We verified that the analytical results proposed in this letter are well-matched with the simulation results. Based on the analytical results, we can figure out the fundamental characteristics of the RH-GFMA system and clearly explain the effect of resource collisions on error performance.

## REFERENCES

[1] S. Dang, O. Amin, B. Shihada, and M. S. Alouini, "What should 6G be?," *Nat. Electron.*, vol. 3, no. 1, pp. 20–29, Jan. 2020.

[2] D. C. Nguyen et al., "6G Internet of Things: A comprehensive survey," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 359–383, Jan. 2022.

[3] S. R. Pokhrel, J. Ding, J. Park, O. S. Park, and J. Choi, "Towards enabling critical mMTC: A review of URLLC within mMTC," *IEEE Access*, vol. 8, pp. 131796–131813, 2020.

[4] N. H. Mahmood et al., "Machine type communications: Key drivers and enablers towards the 6G era," *EURASIP J. Wireless Commun. Netw.*, vol. 2021, no. 1, pp. 1–25, Jun. 2021.

[5] J. Kim et al., "Two-step random access for 5G system: Latest trends and challenges," *IEEE Netw.*, vol. 35, no. 1, pp. 273–279, Jan./Feb. 2021.

[6] J. Choi, "NOMA-based random access with multichannel ALOHA," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2736–2743, Dec. 2017.

[7] J. B. Seo, B. C. Jung, and H. Jin, "Performance analysis of NOMA random access," *IEEE Commun. Lett.*, vol. 22, no. 11, pp. 2242–2245, Nov. 2018.

[8] H. S. Jang, B. C. Jung, T. Q. S. Quek, and D. K. Sung, "Resource-hopping-based grant-free multiple access for 6G-enabled massive IoT networks," *IEEE Internet Things J.*, vol. 8, no. 20, pp. 15349–15360, Oct. 2021.

[9] L. Liu et al., "Sparse signal processing for grant-free massive connectivity: A future paradigm for random access protocols in the Internet of Things," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 88–99, Sep. 2018.

[10] M. Shirvanimoghaddam et al., "Short block-length codes for ultra-reliable low latency communications," *IEEE Commun. Mag.*, vol. 57, no. 2, pp. 130–137, Feb. 2019.

[11] "Contribution R1-1611108-evaluation on channel coding candidates for URLLC and mMTC," 3GPP, Reno, NV, USA, document 3GPP Working Group Meeting RAN1 #87, Nov. 2016.

[12] J. S. Yeom, H. S. Jang, K. S. Ko, and B. C. Jung, "BER performance of uplink NOMA with joint maximum-likelihood detector," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 10295–10300, Oct. 2019.

[13] W. Ayoub, A. E. Samhat, F. Nouvel, M. Mroue, and J. -C. Prévotet, "Internet of Mobile Things: Overview of LoRaWAN, DASH7, and NB-IoT in LPWANs standards and supported mobility," *IEEE Commun. Surveys Tuts.*, vol 21, no 2, pp. 1561–1581, 2nd Quart. 2019.

[14] A. J. Viterbi, "Convolutional codes and their performance in communication systems," *IEEE Trans. Commun. Technol.*, vol. COM-19, no. 5, pp. 751–772, Oct. 1971.

[15] J. Conan, "The weight spectra of some short low-rate convolutional codes," *IEEE Trans. Commun.*, vol. COM-32, no. 9, pp. 1050–1053, Sep. 1984.

[16] B. C. Jung and D. K. Sung, "Performance analysis of orthogonal code hopping multiplexing systems with repetition, convolutional, and turbo codes," *IEEE Trans. Veh. Technol.*, vol. 57, no. 2, pp. 932–944, Mar. 2008.

[17] "Technical specification group radio access; study on new radio (NR) access technology (release 16)," 3GPP, Sophia Antipolis, France, Rep. TR 38.912, V.16.0.0, Jul. 2020.

[18] H. Ji et al., "Ultra-reliable and low-latency communications in 5G downlink: Physical layer aspects," *IEEE Wireless Commun.*, vol. 25, no. 3, pp. 124–130, Jun. 2018.