Minimizing Age-of-Information in HARQ-CC Aided NOMA Systems

Shaohua Wu, *Member, IEEE*, Zhihong Deng, Aimin Li, *Graduate Student Member, IEEE*, Jian Jiao, *Member, IEEE*, Ning Zhang, *Senior Member, IEEE*, and Qinyu Zhang, *Senior Member, IEEE*

***Abstract*— In this paper, we investigate the timeliness performance of a downlink wireless communication system with non-orthogonal multiple access (NOMA). The timeliness of the system is characterized by Age of Information (AoI). To efficiently utilize the time-frequency resource and achieve a tradeoff between timeliness and reliability, we propose an adaptive transmission policy under hybrid automatic repeat request with chase combining (HARQ-CC) aided NOMA systems. In particular, the BS can adaptively adjust the power allocation and decide whether to transmit old or new packets to users in the NOMA system, based on the current AoI status and the positive/negative acknowledgement (ACK/NACK) feedback signal. We first analyze the BLER under such adaptive systems, and then formulate an AoI minimization problem based on the derived BLER. By transforming the objective function to a Markov Decision Process (MDP) problem, an optimal policy is obtained to minimize the average AoI of the system. Considering the high complexity of the MDP, we further divise an alternative near-optimal policy based on Lyapunov Drift function. Furthermore, we consider the fairness of users and propose a greedy policy to minimize the maximal expected AoI of users. Based on extensive simulations, it has been found that NOMA can outperform OMA on both an overall and a user-level basis when operating with adaptive retransmission and power allocation strategies.**

***Index Terms*— Age of information, NOMA, HARQ-CC, finite blocklength, MDP, Lyapunov drift function, power allocation.**

Manuscript received 1 July 2021; revised 19 January 2022 and 13 June 2022; accepted 20 August 2022. Date of publication 2 September 2022; date of current version 13 February 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 61871147 and Grant 62071141, in part by the Shenzhen Science and Technology Program under Grant GXWD20201230155427003-20200730122528002, and in part by the Major Key Project of Peng Cheng Laboratory (PCL) under Grant PCL2021A03-1. The associate editor coordinating the review of this article and approving it for publication was M. C. Gursoy. *(Corresponding author: Shaohua Wu.)*

Shaohua Wu, Jian Jiao, and Qinyu Zhang are with the Department of Electronics and Information Engineering, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China, and also with the Peng Cheng Laboratory, Shenzhen 518055, China (e-mail: hitwush@hit.edu.cn; jiaojian@hit.edu.cn; zqy@hit.edu.cn).

Zhihong Deng and Aimin Li are with the Department of Electronics and Information Engineering, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China (e-mail: 19s152065@stu.hit.edu.cn; liaimin@stu.hit.edu.cn).

Ning Zhang is with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada (e-mail: ning.zhang@uwindsor.ca).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2022.3201115.

Digital Object Identifier 10.1109/TWC.2022.3201115

I. Introduction

A. Motivation

WITH penetration of the 5th Generation Mobile Communication Technology (5G) in various fields, it expects to accommodate the Internet of Everything [1]. In the Internet of things (IoT) applications focusing on real-time monitoring, the control center needs to know the status of equipments in real time, and sends devices timely messages or control commands, in which the timeliness of information plays an important role. However, traditional performance metrics, such as throughput and delay, are insufficient to capture the information timeliness at the receiver. For this purpose, S. Kaul *et. al* propose Age of Information (AoI) as a new metric to characterize the freshness of information from the perspective of the receiver [2]. AoI can be defined as the time elapsed since the most recently successfully received status update was generated at the destination. Meanwhile, it is found that the information age cannot be minimized by maximizing throughput in [3]. Further discussions on minimizing AoI can be referred to [4], [5] in which the control policies by preempting [4] or dropping [5] the old packet in queues are devised.

As devices share access medium in wireless networks, the research on the timeliness of different access strategies for multi-source systems has been widely studied. Because of the superior spectral efficient compared to orthogonal multiple access (OMA), non-orthogonal multiple access (NOMA) is considered as a promising access technique for the future network [6]. Some works study the performance comparison in terms of throughput [7] and delay [8] between NOMA and OMA, respectively. These works demonstrate that NOMA can outperform OMA under the Signal to Noise Ratio (SNR) constraint, through allowing multiple users to transmit information by using overlay technology in the same resource block. Motivated by the potential in improving the traditional metrics, the timeliness performance of NOMA can be found in [9], [10], and [11]. All these researches reveal that when the channel condition cannot guarantee reliable transmissions, the AoI performance of NOMA may be worse than OMA, and thus the advantage of NOMA in spectrum efficiency cannot be embodied.

In order to improve the transmission reliability, Hybrid Automatic Repeat reQuest (HARQ) which integrates forward error correction code and automatic retransmission request was proposed in [12]. HARQ uses time diversity to ensure reliable transmission, but it will cause information aging. However, as shown in [13] and [14], by appropriately choosing the maximal number of retransmissions and the time to retransmit, HARQ can improve the AoI performance. Moreover, in the multi-source communication system, there is a trend to combine HARQ with NOMA to achieve the performance tradeoff at the system level. The literature on enhancing the throughput and reducing the delay of NOMA systems through HARQ can be found in [15] and [16]. In terms of system timeliness, HARQ can improve the transmission reliability of NOMA, so that it can take full advantage of the high spectral efficiency of NOMA to mitigate the information aging caused by retransmissions. Nevertheless, the research on timeliness analysis and optimization of HARQ aided NOMA systems is still relatively scarce.

Moreover, the far user is inferior to the near user in terms of resource competition due to the weak channel condition, which will thus result in a non-negligible performance gap among users. Thus, to prevent the outage of the far user caused by outdated information, the user fairness is a noteworthy issue in NOMA. The works that devote to guaranteeing the fairness of NOMA are investigated in [17], [18], and [19], which reveal that NOMA can achieve better rate fairness performance than OMA by using maximin and proportional fairness power allocation schemes. Thus, to further exploit the potential timeliness of HARQ aided NOMA systems, the analysis on the AoI performance of individual users is necessary.

B. Related Works

The related works on minimizing AoI by various algorithms can be found in [14], [20], [21], and [22]. As the action chosen by the system will result in different AoI evolutions, the MDP problem is formulated to obtain the optimal transmission policy to minimize AoI in [14], [20], and [21]. The scheduling algorithm based on Lyapunov Drift function [22] is proposed to optimize the AoI of the system through appropriate allocation of channel resources. In [23], Pan *et al.* compared the AoI performance in TDMA and FDMA under a generate-at-will model, which extended the analysis to the physical layer. For AoI analysis of NOMA, in [10] and [24], the authors construct a hybrid NOMA/OMA scheme to optimize the AoI performance of multi-user communication systems. In [10], the authors solve the MDP problem to obtain the optimal power allocation strategy in the NOMA system. At the same time, in order to deal with the disadvantage on computational complexity of MDP, a suboptimal scheme based on maximum weight strategy is proposed. Similarly, in [24], an opportunistic NOMA/OMA scheme is used to minimize the average AoI of users with diverse QoS requirements based on the Lyapunov framework. Furthermore, inspired by the SIC in NOMA, the authors in [25] investigate the AoI performance of the slotted ALOHA. In [15], [16], and [26], the authors discussed the effect of HARQ with Chase combining (CC) on the performance of NOMA. Among them, considering the application scenario of uRLLC, [16] extends the work of [26] to the field of finite blocklength, and shows that NOMA has lower delay than OMA under the reliability constraint. In [15], by designing the reasonable user power allocation scheme, the HARQ-NOMA achieves higher throughput than HARQ-OMA under the same average received power constraint at the BS. The works on power allocation for proportional and max-min fairness can be found in [18] and [19]. However, there is little research that analyzed and optimized HARQ-CC aided NOMA systems from the perspective of AoI.

Motivated by above works, this paper contributes new results in the area of AoI analysis, wherein the AoI of HARQ-CC aided NOMA system are analyzed and optimized. We show that the introduction of HARQ-CC can efficiently improve the AoI performance of NOMA systems, particularly when channel conditions are unsatisfactory. Meanwhile, we propose an adaptive transmission policy, including power allocation and retransmission decision, to further improve the AoI performance of HARQ-CC aided NOMA systems. Specifically, we formulate a MDP problem to obtain the optimal policy to achieve the excellent AoI performance. Then, a suboptimal policy based on Lyapunov Drift function is proposed to reduce the computational complexity of MDP. Moreover, a greedy policy on minimizing the maximal expected AoI of users is conducted to analyze how HARQ-CC can help NOMA to equalize the AoI fairness between users.

C. Contributions

The main contributions of this paper are summarized as follows:

- We consider a more flexible scenario where the power allocations and retransmission decisions are made adaptively in each time slot. While in existing works, the BLER analysis were conducted with predefined fixed transmission scheme. The explicit approximate average blocklength error ratio (BLER) of users in the FBL regime are derived under our considered adaptive HARQ-CC systems, which extends the results in [16] and [26].
- Based on the derived average BLER, we design an optimal transmission policy for the adaptive HARQ-CC aided NOMA system by formulating and solving an MDP problem. A stationary optimal policy is proved to exist to achieve the optimal performance. To circumvent the high computational complexity of MDP, we further propose a suboptimal policy by leveraging the Lyapunov Drift function. The simulation results illustrate that the suboptimal policy can achieve near-optimal performance.
- We propose a policy to minimize the user's maximal expected AoI to analyze the performance of individual users' AoI, which can achieve close-performance with OMA on the user fairness. Through the comprehensive analysis on both system and user levels, it illustrates that the HARQ-CC aided NOMA systems can efficiently use time-frequency resources to achieve the tradeoff between transmission reliability and timeliness.

The remainder of this paper is organized as follows. In Section II, we introduce the system model of HARQ-CC aided NOMA systems including the models of AoI evolution and transmission signal. Section III gives the derivative process of average BLERs of individual users and based on the average BLER and retransmission, the problems about AoI optimization is established. For minimizing the expected weighted-sum AoI of both users, an MDP problem is formulated and the suboptimal algorithm based on Lyapunov Drift function is proposed in Section IV. Section V focuses on analyzing the performance of individual users' AoI and a policy on minimizing the user's maximal expected AoI is presented. Simulation results are presented in Section VI. Finally, Section VII concludes the paper.

II. System Model and Preliminaries

A. Transmission Model of HARQ-CC Aided NOMA System

Consider a downlink NOMA system with one Base Station (BS) and N users ($u_i, i \in \{1, \ldots, N\}$) served over a shared channel. The distances from allocated users to the BS are ordered as $d_1 \leq d_2 \leq d_3 \leq \cdots \leq d_N$. We consider a scenario where the BS can only capture the knowledge of stochastic channel state information (CSI). In such a case, denote $h_{i,k} \sim \mathcal{CN}(0, 1)$ as the quasi-static Rayleigh fading of u_i in the k-th transmission, the corresponding channel coefficient between user u_i and BS can be given as

$$\tilde{h}_{i,k} = \sqrt{d_i^{-\eta}} h_{i,k}, \tag{1}$$

where η represents the path loss exponent.

For a downlink NOMA system, the BS generally leverages the superposition coding technique to multiplex the users into power domain. Specifically, the signal received by u_i in the k-th transmission can be denoted as

$$y_{i,k} = \tilde{h}_{i,k} \left(\sum_{i=1}^{N} \sqrt{\alpha_{i,k} P} x_{i,k} \right) + n_{i,k}, \tag{2}$$

where P denotes the total power, $\alpha_{i,k}$ represents the power allocation coefficient of user i in the k-th transmission, such that $\sum_{i=1}^{N} \alpha_{i,k} = 1$, and $n_{i,k}$ is the unit complex additive white Gaussian (AWGN) noise with zero mean and variance $\sigma^2 = 1$.

At the receiver end of NOMA system, SIC is utilized to demultiplex the superimposed signals. The decoding process of SIC performs in an iterative manner, wherein the stronger users u_i with higher channel quality iteratively decode and eliminate the signals from weaker users $\{u_j : j < i\}$, and then decode its own signal. In such a scenario, the allocated power for users always satisfies that $\alpha_{1,k} < \alpha_{2,k} < \cdots < \alpha_{N,k}$, which means that the weaker users are allocated with higher power to enhance the fairness of the served users.

B. User Pairing

In a practical NOMA-based application scenario, user pairing is an extensively employed method to circumvent the high complexity of global recourse scheduling[1]. Generally, paired users are assumed to share a common subchannel in NOMA systems. There are mainly three pairing schemes of NOMA systems, which are summarized as follows:

- *Random schemes* [28]. In such a case, the users are paired randomly, regardless of the channel states of users.
- *Adjacent pairing schemes* [29]. Users having adjacent channel states are paired into some two-elements clusters.
- *Strong-weak pairing schemes* [30]. Such a scheme are proved to an optimal one for NOMA systems. Briefly speaking, user u_i is paired with user u_{N-i+1} in our considered multi-user systems.

With user pairing, the problem of multi-user ($N > 2$) resource scheduling can be reduced to some two-user child problems. As such, we then focus on a two-user NOMA system in the sequel, as shown in Fig. 1.

[1]A two-user NOMA-based communication system, known as Multi-User Superposition Transmission (MUST), has been investigated for the 3GPP Long Term Evolution (LTE) [27].

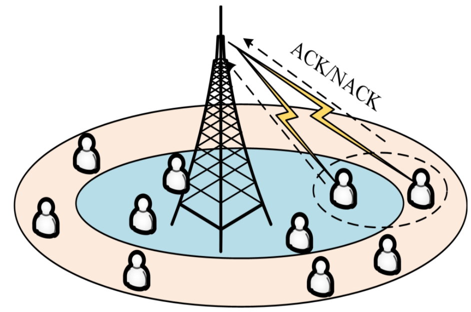


Fig. 1. Downlink NOMA system.

C. AoI Evolution Model of HARQ Aided NOMA

In the user-paired NOMA system, we assume that BS will transmit freshness-critical information to the paired users u_i, $i = \{1, 2\}$ in a time-slotted manner [10]. Without loss of generality, we assume that $d_1 \leq d_2$. The generate-at-will model is adopted in our considered system, wherein the BS can generate packets for both users at the beginning of each time slot. Meanwhile, in order to ensure reliable reception of updates, we combine the HARQ-CC transmission mechanism with NOMA. In HARQ-CC protocol, if u_i successfully decodes its message, it will send a positive acknowledgement (ACK) to the BS, otherwise a negative acknowledgement (NACK) will be fed back. It is assumed that the feedback channel is error-free and delay-free.

We leverage AoI as a metric to measure the timeliness of the information received by users. The instantaneous AoI of u_i at time slot t is defined as $\Delta_i(t) = t - U_i(t)$, where $U_i(t)$ denotes the timestamp that u_i's latest received information generated. An example of AoI evolution of HARQ-CC aided NOMA is shown in Fig. 2. When the user receives a new message from the BS, its AoI will decrease to 1, while if the old message through retransmission is received by u_i, its AoI will decrease to 2, because the old message is generated before 2 time slots. The AoI evolution of u_i between two adjacent time slots can be expressed as

$$\Delta_i(t+1) = \begin{cases} \Delta_i(t) + 1, & \phi_i(t) = 0, \\ \tilde{t}_i, & \phi(t) = \tilde{t}_i, \tilde{t}_i = 1, 2, \cdots, T_{\max} \end{cases} \tag{3}$$

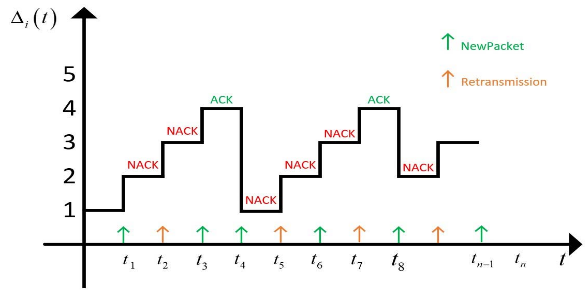


Fig. 2. The AoI evolution of HARQ-CC aided NOMA system.

where $\phi_i(t)$ is associated with the cumulative number of failed decoding rounds of user u_i before current successful decoding time slot t, with $\phi_i(t) = 0$ denoting a decoding failure at current time slot t.

III. Average Blocklength Error Rate and Problem Formulation

The updating process of user's instantaneous AoI is associated with the BLER of each transmission round, which is naturally affected by the power allocation strategy and the retransmission rounds. In this section, we first analyze the interference plus noise ratio (SINR) of each user in each transmission round. With these analytical SINRs in hand, we then derive explicit BLERs in the finite block length (FBL) regime.

A. SINR Analysis

In the k-th transmission round, the weak user u_2 can directly decode its message x_2 through $y_{2,k}$, by treating x_1 as an interference. Therefore, the SINR at u_2 for decoding x_2 is:

$$\gamma_{22}^{(k)} = \alpha_{2,k}\left|\tilde{h}_{2,k}\right|^2 / \left(\alpha_{1,k}\left|\tilde{h}_{2,k}\right|^2 + 1/\rho\right), \tag{4}$$

where $\rho = P/\sigma^2$ denotes the SNR and $\left|\tilde{h}_{2,k}\right|^2$ follows the exponential distribution.

For strong user u_1, the SIC decoder is adopted to decode its information. In particular, u_1 decodes u_2 's message first by treating x_1 as interference, and then decode its own message after eliminating the interference. The SINRs of u_1 decoding x_2 and x_1 are given as:

$$\begin{aligned}\gamma_{12}^{(k)} &= \alpha_{2,k}\left|\tilde{h}_{1,k}\right|^2 / \left(\alpha_{1,k}\left|\tilde{h}_{1,k}\right|^2 + 1/\rho\right)\\ \gamma_{11}^{(k)} &= \alpha_{1,k}\left|\tilde{h}_{1,k}\right|^2 \rho.\end{aligned} \tag{5}$$

In the HARQ-CC aided NOMA system, the user can store the packets which were unsuccessfully decoded and then use the MRC with the retransmission message to improve the transmission reliability. Denote the transmission round of user u_j's message as $\tilde{t_j}$. Utilizing the MRC technique, the cumulative SINR of u_i decoding the message of u_j after $\tilde{t_j}$-th transmission rounds can be denoted as

$$\gamma_{ij}\left(\tilde{t_j}\right) = \sum_{k=1}^{\tilde{t_j}} \gamma_{ij}^{(k)}, \tilde{t_j} = 1, 2, \ldots, T_{\max}. \tag{6}$$

B. FBL Analysis

The BLER in the $\tilde{t_j}$-th transmission round of u_i decoding u_j's message as can be obtained as in [8]:

$$\begin{aligned}\varepsilon_{ij}\left(\tilde{t_j}\right) &\approx Q\left(\frac{C\left(\gamma_{ij}\left(\tilde{t_j}\right)\right) - \frac{N_j}{M}}{\sqrt{V\left(\gamma_{ij}\left(\tilde{t_j}\right)\right)/M}}\right)\\ &= \Psi\left(\gamma_{ij}\left(\tilde{t_j}\right), N_j, M\right),\end{aligned} \tag{7}$$

where N_j is the number of data bits sent by the BS to the user u_j, which are transmitted by using M symbols, $C\left(\gamma_{ij}\left(\tilde{t_j}\right)\right)$ is the channel capacity, and $V\left(\gamma_{ij}\left(\tilde{t_j}\right)\right)$ is the channel dispersion, given as $V\left(\gamma_{ij}\left(\tilde{t_j}\right)\right) = \left(1 - 1\Big/\left(1+\gamma_{ij}\left(\tilde{t_j}\right)\right)^2\right)(\log_2 e)^2$.

The weak user u_2 directly decodes its own message. Therefore, the instantaneous BLER of user u_2 decoding x_2 after $\tilde{t_2}$-th transmission rounds can be approximated as:

$$\varepsilon_2\left(\tilde{t_2}\right) = \varepsilon_{22}\left(\tilde{t_2}\right) \approx \Psi\left(\gamma_{22}\left(\tilde{t_2}\right), N_2, M\right), \tag{8}$$

The close user u_1 needs to implement SIC to decode its own information. By this means, the instantaneous BLER of u_1 is affected by transmission rounds of both u_1 and u_2, given as

$$\varepsilon_1\left(\tilde{t_1}, \tilde{t_2}\right) = \varepsilon_{12}\left(\tilde{t_2}\right) + \left(1 - \varepsilon_{12}\left(\tilde{t_2}\right)\right)\varepsilon_{11}\left(\tilde{t_1}\right), \tag{9}$$

where $\varepsilon_{12}\left(\tilde{t_2}\right)$ and $\varepsilon_{11}\left(\tilde{t_1}\right)$ are given as

$$\begin{aligned}\varepsilon_{12}\left(\tilde{t_2}\right) &\approx \Psi\left(\gamma_{12}\left(\tilde{t_2}\right), N_2, M\right)\\ \varepsilon_{11}\left(\tilde{t_1}\right) &\approx \Psi\left(\gamma_{11}\left(\tilde{t_1}\right), N_1, M\right).\end{aligned} \tag{10}$$

Since we only know the statistical characteristics of the channel, the average BLER is the expectation of $\Psi(\cdot)$ over the distribution of γ_{ij}, given as

$$\bar{\varepsilon}_{ij}\left(\tilde{t_j}\right) \approx \int_0^\infty Q\left(\frac{C(x) - \frac{N_j}{M}}{\sqrt{V(x)/M}}\right) f_{\gamma_{ij}\left(\tilde{t_j}\right)}(x)\mathrm{d}x, \tag{11}$$

where $f_{\gamma_{ij}\left(\tilde{t_j}\right)}(x)$ is the probability density function (PDF) of $\gamma_{ij}\left(\tilde{t_j}\right)$. Due to the complexity of the Q function, it is difficult to obtain the closed-form solution of $\bar{\varepsilon}_{ij}\left(\tilde{t_j}\right)$. Fortunately, Q function can be linearly approximated to overcome the complexity [31], as:

$$Q\left(\frac{C(x) - \frac{N_j}{M}}{\sqrt{V(x)/M}}\right) \approx \begin{cases} 1, & x \le v_{j,M},\\ \frac{1}{2} - \delta_{j,M}\sqrt{M}\left(x - \beta_{j,M}\right), & v_{j,M} < x < \mu_{j,M},\\ 0, & x \ge \mu_{j,M},\end{cases} \tag{12}$$

where $\delta_{j,M} = 1/\sqrt{2\pi\left(2^{2N_j/M} - 1\right)}$, $\beta_{j,M} = 2^{N_j/M} - 1$, $v_{j,M} = \beta_{j,M} - 1/(2\delta_{j,M}\sqrt{M})$ and $\mu_{j,M} = \beta_{j,M} + 1/(2\delta_{j,M}\sqrt{M})$. Thus, the average BLER in (11) can be approximated as in [8]:

$$\bar{\varepsilon}_{ij}\left(\tilde{t_j}\right) \approx \delta_{j,M}\sqrt{M}\int_{v_{j,M}}^{\mu_{j,M}} F_{\gamma_{ij}\left(\tilde{t_j}\right)}(x)\mathrm{d}x, \tag{13}$$

where $F_{\gamma_{ij}}(x)$ represents the cumulative distribution function (CDF) of γ_{ij}. As done in [8], we use Riemann integral ($\int_a^b I(x)\mathrm{d}x = (b-a)I((a+b)/2)$) to approximate (13), which can be expressed as

$$\bar{\varepsilon}_{ij}\left(\tilde{t_j}\right) \approx \delta_{j,M}\sqrt{M}\int_{v_{j,M}}^{\mu_{j,M}} F_{\gamma_{ij}\left(\tilde{t_j}\right)}(x)dx \approx F_{\gamma_{ij}\left(\tilde{t_j}\right)}(\beta_{j,M})$$
$$= \int_0^{\beta_{j,M}} f_{\gamma_{ij}\left(\tilde{t_j}\right)}(x)dx. \tag{14}$$

With (14) in hand, our remaining interest in the sequel is to solve for the distribution of $f_{\gamma_{ij}\left(\tilde{t_j}\right)}(x)$ and then obtain the closed-form expression of $\bar{\varepsilon}_{ij}\left(\tilde{t_j}\right)$.

C. Average BLER With Flexible Power Allocation Strategy

We notice that the existing works in [16] and [26] have derived the average BLER for HARQ-CC aided NOMA systems. However, both of them assume that the power allocations in each retransmission round are fixed, which limits the flexibility of the derived expression. As such, we extend the derivations in [16] and [26], which enables flexible power allocation switching.

1) Average BLER of User u_2: From (5) and (6), we can obtain the PDF of the SINR to decoding u_2's message, given as

$$f_{\gamma_{i2}^{(k)}}(y) = \frac{\alpha_{2,k}}{\lambda_i\rho(\alpha_{2,k}-\alpha_{1,k}y)^2} e^{-\frac{y}{\lambda_i\rho\left(\alpha_{2,k}-\alpha_{1,k}y\right)}}, \tag{15}$$

where $0 < y < \frac{\alpha_{2,k}}{\alpha_{1,k}}$. Note that $\gamma_{ij}\left(\tilde{t_j}\right) = \sum_{k=1}^{\tilde{t_j}} \gamma_{ij}^{(k)}$, the PDF of $\gamma_{ij}\left(\tilde{t_j}\right)$ can be given as the convolutions of $\gamma_{ij}^{(k)}$:

$$f_{\gamma_{ij}\left(\tilde{t_j}\right)}(y) = f_{\gamma_{ij}^{(1)}}(y) * f_{\gamma_{ij}^{(2)}}(y) * \cdots * f_{\gamma_{ij}^{(\tilde{t_j})}}(y). \tag{16}$$

Thus, by applying (15) in (16), we can obtain the following lemma.

Lemma 1: The distribution of $\gamma_{i2}\left(\tilde{t_2}\right)$ can be given by

$$f_{\gamma_{i2}\left(\tilde{t_2}\right)}(y) = \sum_{n_1=1}^{N}\sum_{n_2=1}^{N}\cdots\sum_{n_{\tilde{t_2}}=1}^{N}\left(\prod_{j=1}^{\tilde{t_2}} c_j^{(i)}\mathcal{G}_j^{(i)}\left(a_{n_j}\right)\right)$$
$$\times\frac{\ln(2)}{t}\sum_{k=1}^{2L}\zeta_k e^{-\frac{k\ln(2)}{2y}\sum_{j=1}^{\tilde{t_2}}\theta_j\left(a_{n_j}+1\right)}, \tag{17}$$

where $c_j^{(i)} = \left(2\alpha_{2,j}\theta_j\pi/N\lambda_i\rho\right)\cdot\exp\left(1/\alpha_{2,j}\lambda_i\rho\right)$, $\theta_j = \alpha_{2,j}/\alpha_{1,j}$, $\lambda_i = 1/d_i^{\eta}$, $a_{n_j} = \cos\left(\frac{2n_j-1}{2N}\pi\right)$, ζ_k, $\mathcal{G}_j^{(i)}(x)$ *are defined as*

$$\zeta_k \triangleq (-1)^{L+k}\sum_{\lfloor j=\frac{k+1}{2}\rfloor}^{\min(k,L)}\frac{j\,(L+1)}{(L)!}\binom{L}{j}\binom{2j}{j}\binom{j}{k-j}, \tag{18}$$

and

$$\mathcal{G}_j^{(i)}(x) \triangleq \frac{\sqrt{1-x^2}}{\left(2\alpha_{1,j}-\alpha_{2,j}\theta_j(x+1)\right)^2}$$
$$\times\exp\left(-\frac{2\alpha_{1,j}}{\alpha_{2,j}\lambda_i\rho\left(2\alpha_{1,j}-\alpha_{2,j}\theta_j(x+1)\right)}\right), \tag{19}$$

respectively; N, L are complexity-accuracy trade-off parameters.

Proof: Please refer to Appendix A. □

Then, by adopting (17) into (14), we obtain the following theorem.

Theorem 1: The average BLER of user u_2 in HARQ-CC aided NOMA systems can be approximated as follows:

$$\bar{\varepsilon}_{i2}\left(\tilde{t_2}\right)$$
$$\approx\sum_{n_1=1}^{N}\sum_{n_2=1}^{N}\cdots\sum_{n_{\tilde{t_2}}=1}^{N}\left(\prod_{j=1}^{\tilde{t_2}} c_j^{(i)}\mathcal{G}_j^{(i)}\left(a_{n_j}\right)\right)$$
$$\times\sum_{m=1}^{N}\frac{\pi\ln(2)\sqrt{1-a_m}}{N\beta_{i,M}\left(a_m+1\right)}\sum_{k=1}^{2L}\zeta_k e^{-\frac{k\ln(2)}{\beta_i\left(a_m+1\right)}\left(\sum_{j=1}^{\tilde{t_2}}\theta_j\left(a_{n_j}+1\right)\right)}. \tag{20}$$

Proof: Please refer to Appendix B. □

2) Average BLER of User u_1: The SINR of u_1 decoding x_1 after $\tilde{t_1}$ transmission rounds can be derived as

$$\gamma_{11}\left(\tilde{t_1}\right) = \sum_{k=1}^{\tilde{t_1}}\gamma_{11}^k = \sum_{k=1}^{\tilde{t_1}}\alpha_{1,k}\lambda_1\rho|h_{1,k}|^2. \tag{21}$$

As $|h_{1,k}| \sim \mathcal{CN}(0,1)$, we can obtain that $|h_{1,k}|^2$ is an exponential variable, with $\alpha_{1,k}\lambda_1\rho|h_{1,k}|^2 \sim \mathrm{Exp}\left(1/\left(\alpha_{1,k}\lambda_1\rho\right)\right)$. Denote $\mathcal{Z}_k = 1/\left(\alpha_{1,k}\lambda_1\rho\right)$, the question reduces to solve the distribution of the sum of $\tilde{t}$ independent exponential random variables $X_k \sim \mathrm{Exp}\left(\mathcal{Z}_k\right), k = 1,\ldots,\tilde{t}$.

Lemma 2: (Erlang Distribution) The sum of n i.i.d exponential random variables $Y = \sum_{i=1}^{n} Y_i$, with each random variable $Y_i \sim \mathcal{Z}$, has the Erlang density with parameters $(n,\mathcal{Z})$, given by

$$f_Y(y) = \frac{\mathcal{Z}^n y^{n-1}}{(n-1)!}\exp\left(-\mathcal{Z}y\right). \tag{22}$$

Then, assume that the set of values of $\left(\mathcal{Z}_k\right)_{k=1,\ldots,\tilde{t_1}}$ is $\{q_1,q_2,\ldots,q_\kappa\}$, and define $\mathcal{I}_j \triangleq \{X_k : \mathcal{Z}_k = q_j\}$, we obtain that the sum of the elements in $\mathcal{I}_j$, denoted as $V_j = \sum_{i\in\mathcal{I}_j} i$, follows the Erlang distribution with parameters $\left(\left|\mathcal{I}_j\right|, q_j\right)$. Since $\gamma_{11}\left(\tilde{t_1}\right) = \sum_{k=1}^{\tilde{t_1}} X_k = \sum_{j=1}^{\kappa} V_j$, the question is further simplified to solve the distribution of the sum of κ independent Erlang random variables $V_j \sim \mathrm{Erlang}\left(\left|\mathcal{I}_j\right|, q_j\right)$, $j = 1,\ldots,\kappa$.

Lemma 3: (Sum of Erlang Random Variables) The distribution of $\gamma_{11}\left(\tilde{t_1}\right) = \sum_{j=1}^{\kappa} V_j$ can be given by

$$f_{\gamma_{11}\left(\tilde{t_1}\right)}(y) = \sum_{j=1}^{\kappa} q_j^{|\mathcal{I}_j|}e^{-q_j y}\sum_{n=1}^{|\mathcal{I}_j|}\frac{(-1)^{|\mathcal{I}_j|-n}\,y^{n-1}}{(n-1)!}$$
$$\times\sum_{\substack{\sum_{i=1}^{\kappa} r_i = |\mathcal{I}_j|-n \\ j\neq i,\, r_i\geq 0}}\prod_{\substack{m=1\\ m\neq j}}^{\kappa}\binom{|\mathcal{I}_m|+r_m-1}{r_m}$$
$$\times\frac{q_m^{|\mathcal{I}_m|}}{\left(q_m-q_j\right)^{|\mathcal{I}_m|+r_m}}. \tag{23}$$

Proof: Please refer to Corollary 6.2 in [32] for the detailed proof. □

Applying (23) in (14) results in the average BLER as follows.

Theorem 2: The average BLER of user u_1 in HARQ-CC aided NOMA systems can be approximated as follows:

$$\begin{aligned}\bar{\varepsilon}_{11}\left(\tilde{t_1}\right) \approx & \sum_{j=1}^{\kappa} q_j^{|\mathcal{I}_j|} \sum_{n=1}^{|\mathcal{I}_j|} \frac{(-1)^{|\mathcal{I}_j|-n}\,\Gamma\left(n, q_j\beta_{j,M}\right)}{q_j^n\,(n-1)!} \\ & \times \sum_{\substack{\sum_{i=1}^{\kappa} r_i = |\mathcal{I}_j|-n \\ j\neq i, r_i\geq 0}} \prod_{\substack{m=1\\ m\neq j}}^{\kappa} \binom{|\mathcal{I}_m|+r_m-1}{r_m} \\ & \times \frac{q_m^{|\mathcal{I}_m|}}{(q_m-q_j)^{|\mathcal{I}_m|+r_m}}, \end{aligned} \tag{24}$$

where $\Gamma(n,x) = \int_0^x t^{n-1}e^{-t}dt$ *is the lower incomplete Gamma function.*

Proof: Please refer to Appendix C. □

D. Problem Formulation

In the HARQ-CC aided NOMA system, the user can implement MRC decoding to improve the successful decoding probability. Evidently, the more power the BS allocates to the user, the higher decoding probability the user enjoys. Moreover, the retransmission process performs as a double-edged sword: on the one hand, retransmission implies more redundancy and longer latency, which may result in stale delivery; on the other hand, retransmission may significantly improve the reliability, and thus enhance the timeliness. Therefore, it is imperative for the BS to adaptively adjust the transmission scheme according to the user's AoI. In Fig. 3, we briefly demonstrate how HARQ-CC influences the AoI performance of NOMA systems. With SNR increasing, the AoI performance of always-retransmission decreases, while that of non-retransmission policy gradually conducts better performance than the former one. As such, there is a trade-off between retransmitting old packets and transmitting new packets. Therefore, it is necessary to investigate when and whether to retransmit redundant old packets, in order to minimize the AoI of HARQ-CC aided NOMA systems.

We also notice that the power allocation policy that minimizes the system average AoI will result in unfairness of two users. In this way, a significant performance gap exists between two users in NOMA. Considering the fairness issue, we also formulate a problem (Problem 2) to avoid the case where the far user with weak channel gains cannot get the timely service.

The problem can be formulated as follows:

Problem 1 (Minimizing the Expected Weighted-Sum AoI of Both Users):

$$\arg\min_{\pi} \lim_{T_{slot}\to\infty} \frac{1}{T_{slot}} \sum_{t=1}^{T_{slot}} \sum_{i=1}^{2} \mathbb{E}\left[\omega_i \Delta_i^{\pi}(t)\right]. \tag{25}$$

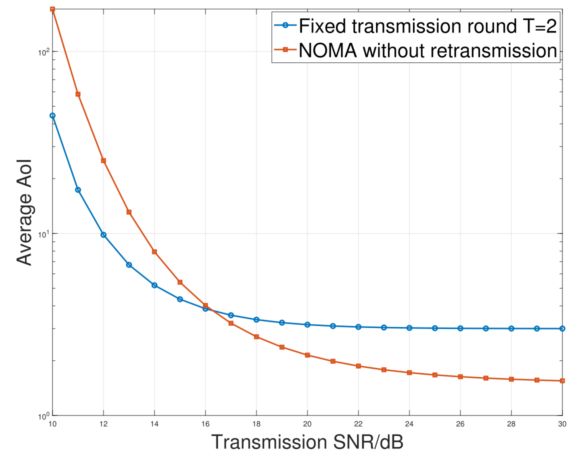


Fig. 3. The AoI performance of NOMA systems with/without HARQ-CC.

Problem 2 (Minimizing the User's Maximal Expected AoI):

$$\arg\min_{\pi} \lim_{T_{slot}\to\infty} \frac{1}{T_{slot}} \sum_{t=1}^{T_{slot}} \max_{i} \left(\mathbb{E}(\Delta_i^{\pi}(t))\right), \tag{26}$$

where π is the policy chosen by the BS including retransmission and power allocation. Problem 1 focuses on optimizing the systemic AoI of HARQ-CC aided NOMA systems. The average AoI is an intuitive manifestation of the system's timeliness. Problem 2 pays more attention to the fairness of different users, which aims to improve the AoI performance from user's perspective. Through the analysis of the above two problems, we can comprehensively illustrate the AoI performance improvement of HARQ-CC aided NOMA systems.

IV. HARQ-CC Aided Power Allocation for Minimizing the Expected Weighted-Sum AoI

Problem 1 can be reformulated as a MDP problem, according to which the optimal stability transmission strategy of the system in each state can be obtained. Meanwhile, we notice that to obtain an optimal policy of the MDP problem will result in high complexity, and thus design a low-complexity algorithm to solve this problem suboptimally.

A. MDP Problem

In this paper, we consider a case study where the maximum retransmission is set as $T_{\max} = 2$ for the implementation simplicity. In such a case, Problem 1 can be transformed to an MDP problem. The MDP problem is characterized by a quaternary tuple $\{\mathcal{S}, \mathcal{A}, \mathcal{P}, r\}$, where $\mathcal{S}, \mathcal{A}, \mathcal{P}$ denotes the state space, action space, and state-action transition probability matrix, respectively.

- **State Space.** The state $s_t \in \mathcal{S}$ of the system at time slot t is characterized by four elements $(a_F, u_F, \Delta_1(t), \Delta_2(t))$, where
  - a_F denotes the power allocation coefficient of the message stored in u_1's buffer, with $a_F \in \{0, 1, 2, \ldots, \lfloor (N_{power} - 1/2 \rfloor\}$. $a_F = 0$ represents that the buffer is empty, while if $a_F \neq 0$, we can know that the power allocation coefficients of the

failed transmission were $\alpha_1 = a_F/N_{power}$ and $\alpha_2 = (N_{power} - a_F)/N_{power}$;

- $u_F \in \{0, 1, 2, 3\}$ denotes the status of the specific user decoding the message transmitted by the BS. If u_i has not completed the decoding process, we set $u_F = i$. If both u_1 and u_2 have not successfully decoded their messages, the element u_F will be turned to 3. Otherwise, u_F will stay 0;
- $\Delta_i(t)$ denote the instantaneous AoI of u_i.

Remark 1: Not all states are required. 1) $a_F = 0$ denotes that there is no message stored in the buffer of both users, which illustrates that the users did not need retransmissions, so we should set $u_F = 0$; 2) $u_F \neq 0$ represents that there is a decoding failure in the last state, thus it can be inferred that $a_F \neq 0$; 3) When $u_F \neq 0$, the corresponding user's AoI should satisfy $\Delta_i(t) > 1$. Because if the user cannot decode the message successfully, its AoI cannot update.

- **Action Space.** The action space $\mathcal{A}$ in the tuple consists of two parts which include retransmission and power allocation. The first part determines whether to transmit new of old packets. In this part, the BS can only retransmit to the users that sent back NACK. Meanwhile, since the maximum transmission round $T_{trans} = 2$, the BS will transmit new message when it has retransmitted a packet, no matter whether the user has decoded successfully or not. The second part determines the power allocation policy. in each time slot, the BS can decide the power allocation to users in a discrete manner $\{a/N_{power} : a = 1, 2, 3, 4\}$.
Remark 2: 1) The BS will implement retransmission only to the user failed to decode the message. 2) If a user received a retransmission packet in the current slot, it has to clean up the buffer, at the end of this slot, whether successfully decoding the message or not. Because the BS will transmit new message in the next time slot, the old message in the buffer cannot improve the success rate of decoding the new information.
- **Reward.** In the time slot t, the reward of each state which is defined as the instantaneous weighted-sum AoI of users, which can be written as $r_t = \omega_1\Delta_1(t) + \omega_2\Delta_2(t)$
- **Transition Probability.** The transition probability from state s_t to s_{t+1} by taking the action a_t can be represented as $P(s_{t+1}|s_t, a_t)$. The transition probabilities under different actions can be represented as
  - The BS transmits new messages to the both users.

$$s_{t+1} = \begin{cases} (0,0,1,1), & \bar{P}_{1,1}\bar{P}_{2,1}, \\ (\hat{a}_F,1,\Delta_1+1,1), & P_{1,1}\bar{P}_{2,1}, \\ (\hat{a}_F,2,1,\Delta_2+1), & \bar{P}_{1,1}P_{2,1}, \\ (\hat{a}_F,3,\Delta_1+1,\Delta_2+1) & P_{1,1}P_{2,1}, \end{cases} \tag{27}$$

where $P_{i,1} = \varepsilon_i(\hat{a}_F)$ and $\bar{P}_{i,1} = 1 - \varepsilon_i(\hat{a}_F)$. The subscript number 1 denotes that there is only one message decoded by u_1.

  - The BS only conducts retransmission to the old message to the close user u_1. The current state is $s_t = (a_F, u_F, \Delta_1(t), \Delta_2(t))$, and $u_F = 1, 3$

$$s_{t+1} = \begin{cases} (0,0,2,1), & \bar{P}_{1,2}\bar{P}_{2,1}, \\ (0,0,\Delta_1+1,1), & P_{1,2}\bar{P}_{2,1}, \\ (\hat{a}_F,2,2,\Delta_2+1), & \bar{P}_{1,2}P_{2,1}, \\ (\hat{a}_F,2,\Delta_1+1,\Delta_2+1,) & P_{1,2}P_{2,1}, \end{cases} \tag{28}$$

where $\hat{a}_F$ is the power allocation coefficient chosen by the BS at state s_t.

  - The BS only retransmits the old message to the far user u_1. The current state is $s_t = (a_F, u_F, \Delta_1(t), \Delta_2(t))$, and $u_F = 2, 3$

$$s_{t+1} = \begin{cases} (0,0,1,2), & \bar{P}_{1,1}\bar{P}_{2,2}, \\ (\hat{a}_F,1,\Delta_1+1,2), & P_{1,1}\bar{P}_{2,2}, \\ (0,0,1,\Delta_2+1), & \bar{P}_{1,1}P_{2,2}, \\ (\hat{a}_F,1,\Delta_1+1,\Delta_2+1), & P_{1,1}P_{2,2}, \end{cases} \tag{29}$$

where $P_{2,2} = \varepsilon_2(\hat{a}_F, a_F)$ denotes the average BLER of u_2 decoding the retransmission messages.

  - The BS only retransmits the old message to the both users.

$$s_{t+1} = \begin{cases} (0,0,2,2), & \bar{P}_{1,2}\bar{P}_{2,2}, \\ (0,0,\Delta_1+1,2), & P_{1,2}\bar{P}_{2,2}, \\ (0,0,2,\Delta_2+1), & \bar{P}_{1,2}P_{2,2}, \\ (0,0,\Delta_1+1,\Delta_2+1), & P_{1,2}P_{2,2}, \end{cases} \tag{30}$$

where because both users have received the retransmission messages, so in the next state the user need clean up their buffers.

We define the set of all feasible policies by Π and a feasible policy $\pi \in \Pi$. Then, given the initial state s_0, the long-term average reward under any feasible policy π can be expressed as

$$C(\pi, s_0) = \lim_{T_{slots}\to\infty} \sup \frac{1}{T_{slots}} \sum_{t=0}^{T_{slots}} \mathbb{E}^{\pi}\left[r(s_t, a_t)|s_0\right]. \tag{31}$$

The optimization of Problem 1 can be rewritten as
Problem 3:

$$\underset{\pi\in\Pi}{\text{minimize}}\, C(\pi, s_0). \tag{32}$$

B. The Optimal Policy

In this subsection, we first illustrate the existence of the optimal stationary policy of Problem 3 and then present some results of the optimal policy, which can reflect that the policy switches with the change of states.

Theorem 3: There is an optimal policy at any state s denoted by $\pi^*(s)$ and the value function $V(s)$ represents the mapping from state s to the reward set: $S \rightarrow R$. The average reward optimality equation can be expressed as

$$J^* + V(s) = \min_{a\in A}\left\{c(s,a) + \mathbb{E}\left[V(s') \mid s, a\right]\right\}, \quad \forall s \in S \tag{33}$$

where J^* is the optimal average reward with $\pi^*(s)$ and s' is the next state from s *by taking the action* a*.*

Proof: Please refer to Appendix D. □

Based on the stationary optimal policy, the boundary state of retransmission of both users can be derived.

Lemma 4: The retransmission boundary of Δ_1 is determinded by the average BLERs of both users and Δ_2

Proof: We first defined the state-action value function $Q_i(s,a)$ at iteration i and the state value function $V_{i+1}(s)$ at iteration $i+1$ as

$$\begin{aligned} Q_i(s,a) &= c(s,a) + \mathbb{E}\left[V_i(s') \mid s,a\right] \\ V_{i+1}(s) &= \min_{a\in A} Q_i(s,a). \end{aligned} \tag{34}$$

When $a_F \neq 0, u_F = 0$, we denote $a=0$ as transmiting new message to u_1 and $a=1$ as retransmission. Thus, we can derive that

$$\begin{aligned} &Q_i(a_F,1,\Delta_1,\Delta_2,1) - Q_i(a_F,1,\Delta_1,\Delta_2,0) \\ &\quad = 1-\varepsilon_1 + (\varepsilon_1-\tilde{\varepsilon}_1)\Delta_1 + (\varepsilon_2-\tilde{\varepsilon}_2)\Delta_2, \end{aligned} \tag{35}$$

where ε_1 and $\tilde{\varepsilon}_1$ denote the average BLER of u_1 decoding the retransmission message and the new message, respectively, and $\varepsilon_1 \le \tilde{\varepsilon}_1$. Thus, when $\Delta_1 \le \frac{1-\varepsilon_1+(\varepsilon_2-\tilde{\varepsilon}_2)\Delta_2}{\tilde{\varepsilon}_1-\varepsilon_1}$, the BS will transmite new messages to u_1, even if u_1 has failed to decode the message in current state. It is easy to extend the proof to the retransmission boundary of u_2. □

C. Suboptimal Policy Based on Lyapunov Drift Function

Solving the MDP problem requires storing the whole state space and state transition probability matrix in the memory, which will cast a heavy memory footprints and time cost. Therefore, in this subsection, a suboptimal policy is proposed to deal with Problem 1, which is inspired by the Max-weight policy in [22]. In view of the complexity of the adaptive HARQ-CC aided NOMA system, the closed-form expression of the system average AoI is difficult to obtain. Thus, we build a Markov chain to calculate the system average AoI. Based on the transition probability illustrated in the MDP problem, we formulate the Lyapunov Drift function to represent the expected increment of AoI in next state by taking action π.

$$\Phi_\pi(\vec{\Delta}(t)) = \mathbb{E}_\pi[L(\vec{\Delta}(t+1)) - L(\vec{\Delta}(t)) \mid \vec{\Delta}(t)], \tag{36}$$

where $L(\vec{\Delta}(t)) = \frac{1}{N_u}\sum_{i=1}^{N_u}\omega_i\Delta_i(t)$ denotes the instantaneous AoI of the system and $N_u = 2$ in the considered model. $L(\vec{\Delta}(t+1))$ denotes the expected AoI of the next stage based on state s, which can be expressed as

The current state $s=(0,0,\Delta_1,\Delta_2)$:

$$L(\vec{\Delta}(t+1)) = \omega_1\Delta_1 + \omega_2\Delta_2 + 1. \tag{37}$$

The current state $s=(a_F,1,\Delta_1,\Delta_2)$:

$$L(\vec{\Delta}(t+1)) = \begin{cases} \omega_1(\varepsilon_1)_1\Delta_1 + \omega_2(\varepsilon_2)_1\Delta_2 + 1, \\ \omega_1(\varepsilon_1)_2(\Delta_1-1) + \omega_2(\varepsilon_2)_1\Delta_2 + 1, \end{cases} \tag{38}$$

The current state $s=(a_F,2,\Delta_1,\Delta_2)$:

$$L(\vec{\Delta}(t+1)) = \begin{cases} \omega_1(\varepsilon_1)_1\Delta_1 + \omega_2(\varepsilon_2)_1\Delta_2 + 1, \\ \omega_1(\varepsilon_1)_1\Delta_1 + \omega_2(\varepsilon_2)_2(\Delta_2-1) + 1. \end{cases} \tag{39}$$

The current state $s=(a_F,3,\Delta_1,\Delta_2)$:

$$\begin{aligned} &L(\vec{\Delta}(t+1)) \\ &\quad = \begin{cases} \omega_1(\varepsilon_1)_1\Delta_1 + \omega_2(\varepsilon_2)_1\Delta_2 + 1, \\ \omega_1(\varepsilon_1)_2(\Delta_1-1) + \omega_2(\varepsilon_2)_1\Delta_2 + 1, \\ \omega_1(\varepsilon_1)_1\Delta_1 + \omega_2(\varepsilon_2)_2(\Delta_2-1) + 1, \\ \omega_1(\varepsilon_1)_2(\Delta_1-1) + \omega_2(\varepsilon_2)_2(\Delta_2-1) + 1. \end{cases} \end{aligned} \tag{40}$$

The forms of Lyapunov Drift function under different actions can be represented as

a) BS transmits new messages to both user

$$\begin{aligned} \Phi(\vec{\Delta}(t)) = &-w_1\left(1-(\varepsilon_1(\hat{a}))_1\right)\Delta_1(t) \\ &-w_2\left(1-(\varepsilon_2(\hat{a}))_1\right)\Delta_2(t) + 1. \end{aligned} \tag{41}$$

b) BS transmits a new message to u_1 and retransmits the old message to u_2

$$\begin{aligned} \Phi(\vec{\Delta}(t)) = &-w_1\left(1-(\varepsilon_1(a_F,\hat{a}))_2\right)(\Delta_1(t)-1) \\ &-w_2\left(1-(\varepsilon_2(\hat{a}))_1\right)\Delta_2(t) + 1. \end{aligned} \tag{42}$$

c) BS transmits a new message to u_2 and retransmits the old message to u_1

$$\begin{aligned} \Phi(\vec{\Delta}(t)) = &-w_1\left(1-(\varepsilon_1(\hat{a}))_1\right)\Delta_1(t) \\ &-w_2\left(1-(\varepsilon_2(a_F,\hat{a}))_2\right)(\Delta_2(t)-1) + 1. \end{aligned} \tag{43}$$

d) BS retransmits old massages to both users

$$\begin{aligned} \Phi(\vec{\Delta}(t)) = &-w_1\left(1-(\varepsilon_1(a_F,\hat{a}))_2\right)(\Delta_1(t)-1) \\ &-w_2\left(1-(\varepsilon_2(a_F,\hat{a}))_2\right)(\Delta_2(t)-1) + 1. \end{aligned} \tag{44}$$

Then, the action of the power allocation coefficient in the next transmission round can be given by

$$\pi^*(s_t) = \arg\min_{\hat{a}} \Phi(\vec{\Delta}(t)), \tag{45}$$

where by minimizing $\Phi(\vec{\Delta}(t))$, it can be guaranteed that there are always having the minimal AoI increments in the next state, so that an asymptotically suboptimal policy can be obtained. Moreover, the suboptimal policy only needs to consider the current state, which can achieve lower complexity.

V. HARQ-CC Aided Power Allocation for Minimizing the Users' Maximal Expected AoI

For the purpose of optimizing the system AoI of NOMA, the BS will give priority to the updates of the close user which has better channel conditions. Howerver, due to the power allocation and interference, the AoI of the far user will be deteriorated. In the real-time monitoring system, it not only needs to ensure that the AoI of the system is at a low level, but also need to prevent the outage of users caused by unstable timeliness.

By solving Problem 2, we can investigate the fairness between two users in the HARQ-CC aided NOMA systems.

In Problem 2, we aim to obtain the transmission poloicy to minimize the maximal expected AoI of both users, so that the far user will not be outage because of the bad timeliness. Based on the Lyapunov Drift function, a greedy policy is used to solve Problem 2. The optimal objective can be expressed as

$$\Phi\left(s'\right) = \max\left(\mathbb{E}\left(\Delta_1(s')\right), \mathbb{E}\left(\Delta_2(s')\right)\right), \tag{46}$$

where $\mathbb{E}\left(\Delta_1(s')\right)$ is the expected AoI in the next stage of current state s and be represented by (37)-(40). Then, we need to find the policy $\pi^*\left(s\right)$ to minimize the maximal user's AoI, which should satisfy

$$\pi^*\left(s\right) = \arg\min_{\pi\in A} \Phi\left(s'\right). \tag{47}$$

where the policy set A is the same with that in section IV.

For the optimal function (46), it only focuses on the user with maximal expected AoI, so we should select the policy which satisfies (46) and subject to minimize the expected AoI of the other user. We denote the maximal/minimal expected AoI of users in current state as $\Delta_{\sup}\left(s\right)/\Delta_{\inf}\left(s\right)$, the policy about the type of messages as A_{type} and the policy about power allocations as A_{power}. There is a Corollary and the Algorithm 1 gives the algorithm process of the greedy policy.

Corollary 1: Based on (37)-(40), for a state with $a_F \neq 0, u_F \neq 0$, there have multiple kinds of policy A_{type} can be chosen by the BS to transmit the messages. For the user with maximal expected AoI $\Delta_{\sup}\left(s\right)$, there exist an optimal power allocation policy to minimize its AoI, and a further condition should be satisfied as

$$\pi^*\left(s\right) = \arg\min_{\pi\in A_{type}} \mathbb{E}\left(\Delta_{\inf}(s')\right). \tag{48}$$

Through the Min-Max algorithm, the policy focuses on the fairness can be derived and Problem 2 is transferred to a definite Markov chain. The transition matrix of the corresponding Markov chain can be constructed. Then, the stationary probability of each state can be derived, so that the average AoI under the Min-Max policy can be obtained.

Algorithm 1 Min-Max Greedy Policy Algorithm

Input: State space S, numbers of states n, policy set A and transition probability matrix $\mathbf{P}$.
Output: optimal policy set $\pi^*\left(s\right)$

$i = 0$, $\Phi\left(s\right) = 0$
while $i < n$ **do**
  $\left[a_F, u_F, \Delta_1, \Delta_2\right] = \mathrm{decode}\left(s_i\right)$
  if $a_F = u_F = 0$ **then**
    $\pi^*\left(s_i\right) = \arg\min_{\pi\in A} \Phi\left(s'\right), \text{s.t.} \arg\min_{\pi\in A} \mathbb{E}\left(\Delta_{\inf}(s')\right)$
  else
    $L\left(N_{A_{type}}, N_{A_{power}}\right) = \arg\min_{\pi\in A} \Phi\left(s'\right)$
    $\pi^*\left(s_i\right) = \arg\min_{\pi\in A_{type}} \mathbb{E}\left(\Delta_{\inf}(s')\right)$
  $i = i + 1$
return $\pi^*\left(s\right)$

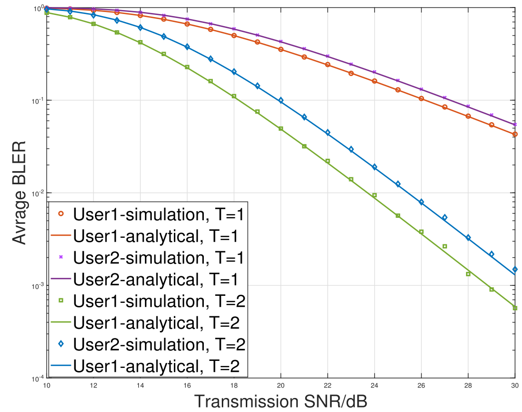


Fig. 4. The average BLER versus transmission SNR with different transmission rounds, where the power allocation coefficients to both users are $\alpha_1 = 0.2, \alpha_1 = 0.8$.

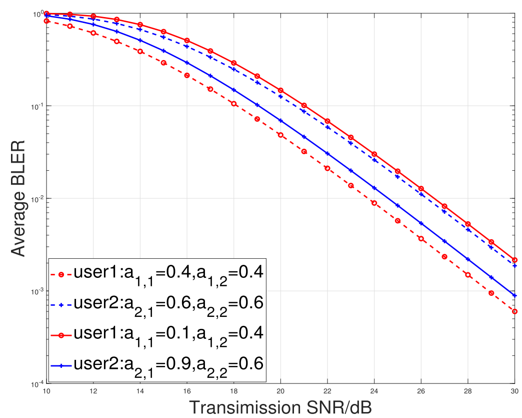


Fig. 5. The average BLER of users with different power allocation coefficients with maximum transmission round $T_{\max} = 2$.

VI. Simulation Results

In this section, we first present the performance in terms of the average BLER of HARQ-CC aided NOMA systems and analyze the influence on transmission reliability caused by retransmissions and power allocations. Then the policy comparison of MDP and Lyapunov Drift function is exhibited. Lastly, the AoI performance is demonstrated. We analyze the improvement of timeliness of NOMA system by HARQ-CC transmission from two dimensions, inluding system average AoI and individual AoI of each user.

A. Average BLER

This subsection presents the BLER performance. We first conduct Monte Carlo simulation to verify the approximations accuracy of the transmission BLER of HARQ-CC aided NOMA derived in Section III. To make sure the numerical accuracy, in the simulation, the complexity-accuracy parameters in (20) are set as $N = 50$ and $L = 20$. The relative distance of the user from the BS is set as $d_1 = 3$ and $d_2 = 7$, respectively. The path loss exponent $\eta = 2$ and $N_{power} = 10$.

Fig. 4 shows the average BLERs versus the transmission SNR with different transmission rounds. We set the information bits $N_1 = N_2 = 160$, and the number of symbols

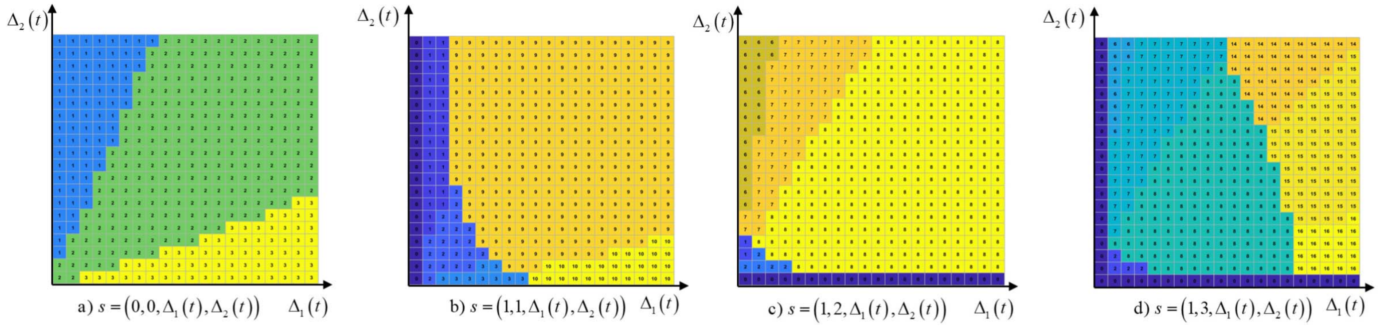


Fig. 6. The optimal policy by solving MDP problem.

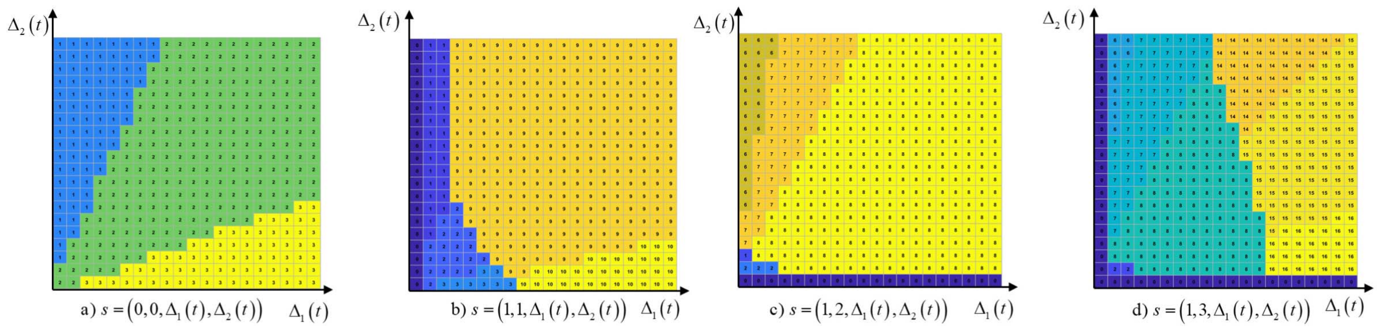


Fig. 7. The suboptimal policy by Lyapunov Drift.

$M = 200$. The closed-form approximations derived by (20) and (24) can match with the Monte Carlo simulation results. With the increasing number of transmission rounds, the average BLER decreases under a given transmission SNR. It illustrates that HARQ-CC can improve the decoding reliability of both users in NOMA systems and because u_1 has a better channel condition, its average BLER is smaller than u_2.

Fig. 5 plots the average BLER of both users under different power allocation coefficients. It presents that when the BS allocates more power to a specific user, the decoding reliability will be increased. For the close user u_1, if it has failed to decode the message with power $p_{1,1} = 0.4P$, the user will have a higher probability to successfully decode the retransmission message with $p_{1,2} = 0.4P$ than with $0.1P$. The analysis of average BLER of users decoding messages with different power allocations can make the BS adaptively switch the policy according to users' current AoI, which is more suitable for practical application.

B. The Optimal and Suboptimal Policy

This subsection is aimed to compare the policies derived by MDP and Lyapunov Drift function. For the MDP problem proposed in (32), we use relative value iteration (RVI) to obtain the optimal policy. In order to get better AoI convergence effect, the relative distance of the user from the BS is set as $d_1 = 2$ and $d_2 = 5$ to ensure that the instantaneous AoI cannot exceed the threshold. Thus, we can set the truncated finite AoI states $\Delta_i \leq 20$ to approximate the countable infinite state space [10]. Moreover, for the sake of ensuring that the users' AoI is within the given threshold, the transmission SNR $\rho = 19$dB. An action dictionary is formed to recognize the policy, $D_s = 4 \cdot T_F + a_s$, where $T_F = \{0, 1, 2, 3\}$ represents retransmission to none user, the far user u_2, the close user u_1 and both users, and $a_s = \{1, \ldots, 4\}$ is the power allocation coefficient to the close user u_1.

Fig. 6 presents the optimal policy by solving the MDP problem. It aims to show the policy switches according to the current state. a) The current $s = (0, 0, \Delta_1(t), \Delta_2(t))$, so the BS can only transmit new messages to both users, where the $T_F = 0$. We can find that with $\Delta_i(t)$, increasing the BS will allocate more power to u_i; b) The current $s = (1, 1, \Delta_1(t), \Delta_2(t))$ and $T_F = 0, 2$, where the u_1 has failed to decode the message, so the BS can retransmit the old packet to u_1. It presents that when the AoI of u_1 is low, the BS will not decide to implement retransmissions; c) If only u_2 needs retransmission, as the AoI increasing, the BS will allocate more power to retransmit the message of u_2; d) The current $s = (1, 3, \Delta_1(t), \Delta_2(t))$, it means that both user feedback NACK to the BS. However, the BS prefers to allocate more power and retransmits messages to the user with high AoI.

The suboptimal policy derived by Lyapunov Drift function is depicted in Fig. 7. We can find that the suboptimal policy is similar to the optimal policy which shows the feasibility of the suboptimal algorithm.

C. The Average AoI Under Different Transmission Policies

The performance of the average AoI of the whole system and individual users by solving Problem 1 and Problem 2 is analyzed respectively, in this subsection. Fig. 8 aims to compare the AoI performance of different communication systems

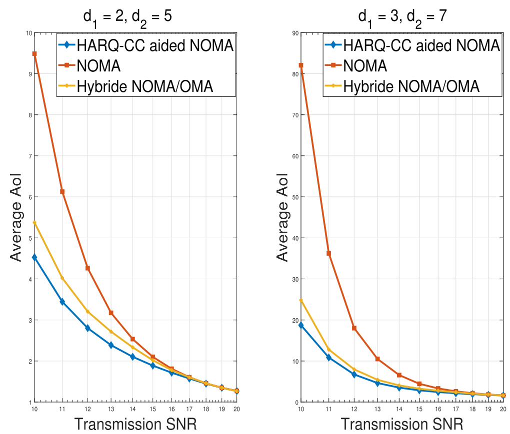


Fig. 8. The performance comparison of different access schemes under optimal policy.

with optimal policy by solving the MDP problem in two cases 1) $d_1 = 2, d_2 = 5$; 2) $d_1 = 3, d_2 = 7$. The power allocation ratio is set as $N_{power} = 10$ and we apply RVI on truncated finite states $(\Delta_i \leq 300)$ to approximate the countable infinite state space. The simulation results of all policies are generated by 10^6 time slots ergodic. According to the optimal policy, we can find that the HARQ-CC aided NOMA can always achieve the best performance in the considered three schemes, including Hybrid NOMA/OMA and NOMA systems.

Fig. 9 illustrates the suboptimal performance of the weighted sum of the expected AoI of the access policies in Fig. 8. The policy based on Lyapunov Drift function can achieve the best performance with lower computational complexity. Combining with Fig.8 and Fig. 9, we can find that, when SNR is low, because HARQ-CC is combined with NOMA, the reliability of NOMA is improved, which brings positive benefits to the AoI performance. The adaptive HARQ-CC aided NOMA that makes a trade-off between reliability and timeliness even obviously outperform the adaptive NOMA/OMA systems. As SNR increases, the performance curves of the three strategies are gradually getting closer. This is because the reliability of each transmission is at a high level. Retransmitting the old packets can only deteriorate the timeliness of the system, so all strategies will choose a single round NOMA transmission.

Fig. 10 depicts the long-term average AoI of the HARQ-CC aided NOMA system with optimal and suboptimal policy. One can see that the suboptimal policy can achieve near-optimal performance, especially, when the distance between the BS and users is small. Moreover, the suboptimal polic based on Lyapunov Drift function can be seen as an online policy, because the transmission only depends on the current state. The optimal policy is more like an offline policy, which can look up the action dictionary and choose the long-term optimal action based on the current state. The similar performance of optimal and suboptimal policies on average AoI can provide the ability to build a hybrid system according to the reality.

Fig. 11 and Fig. 12 respectively show the AoI performance of systemic AoI and individual user's AoI. In Fig. 11, the average AoI under the policy by minimizing the maximal

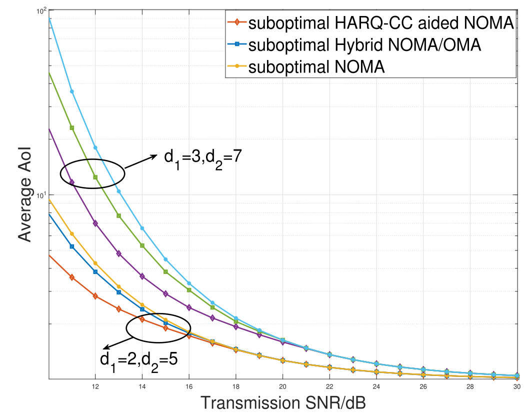


Fig. 9. The performance comparison of different access schemes under suboptimal policy.

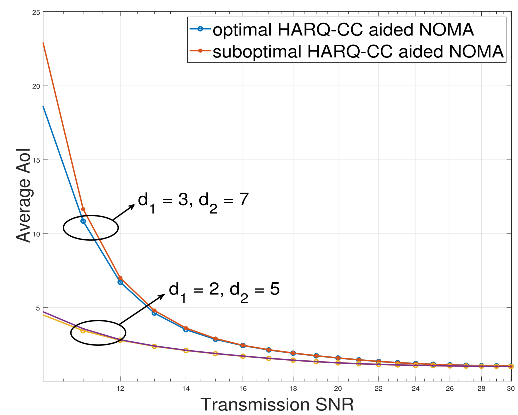


Fig. 10. The average AoI performance of MDP and Lyapunov against transmission SNR.

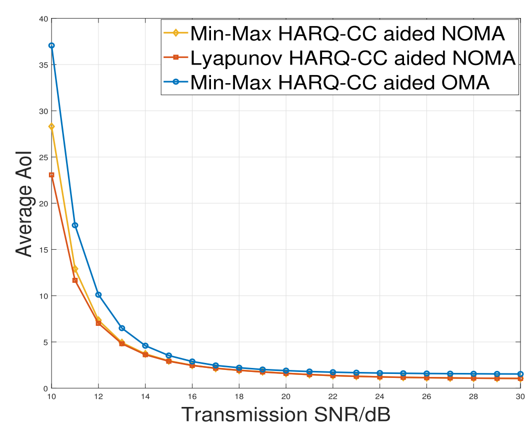


Fig. 11. The comparison of systemic AoI performance between Min-Max policy and suboptimal policy with different access scheme.

expected AoI of both users can achieve near performance to the suboptimal HARQ-CC aided NOMA systems and can outperform the HARQ-CC aided OMA systems. That is because OMA can only serve one user in one transmission, so the other user can access the channel until the user in service completing the process. This disadvantage is magnified when users without good channel conditions, in which the user has to wait for a long time before it can receive the update from the BS. The HARQ-CC aided NOMA can sufficiently improve

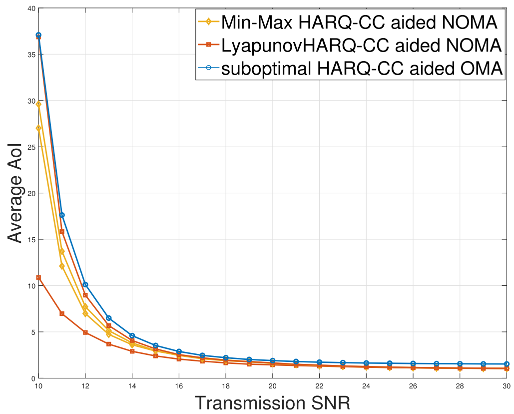


Fig. 12. The comparison of the individual user's AoI performance between Min-Max policy and suboptimal policy with different access scheme.

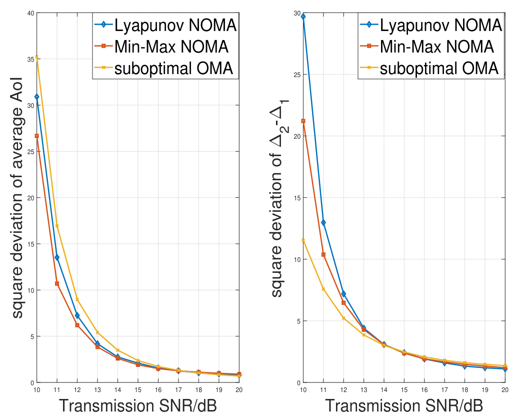


Fig. 13. The comparison of the individual user's AoI performance between Min-Max policy and suboptimal policy with different access scheme.

the reliability of NOMA and the characteristic of serving multiple users in one source block can solve the problem of the information aging cased by retransmission in HARQ-CC.

Fig. 12 depicts that through the Min-Max policy, the AoI gap between two users will become narrow. HARQ-CC can optimize both users' AoI in NOMA and achieve a better performance than OMA. From Fig. 13, the square deviation of AoI is used to illustrate the variation intensity of AoI evolution. We can find that the Min-Max policy has the highest AoI stability of the system. For the AoI gap between two users in the system, OMA, owing to serve only one user by allocating all power to transmit the message, has a smallest gap between two users. However, the HARQ-CC aided NOMA can guarantee the systemic and individual user's AoI in a stable state, simultaneously.

VII. Conclusion and Future Works

In this paper, we analyze the AoI performance of a two-user HARQ-CC aided NOMA system, wherein the BS can adaptively adjust the power allocation strategy and the retransmission policy to minimize the AoI performance of users. Firstly, we extend the existing work on analyzing BLER of HARQ-CC aided NOMA systems. Instead of focusing on the BLER analysis under a predefined transmission strategy, such as power allocations and retransmission decisions, our derived BLERs enable more flexible transmission actions and thus may provide more valuable theoretical guideline for NOMO-based system design. Then, based on the derived BLERs, we formulate an AoI minimization problem, wherein the BS can adaptively adjust the transmission strategy to ensure timely information delivery. We then showcase an example of the optimization process where the maximum retransmission is set as 2, and solve the problem by utilizing MDP method and Lyapunov Drift function to obtain the optimal policy and suboptimal policy, respectively. At last, we further consider the issue of fairness of users. In this regard, we seek to maintain a balance between the AoI performance of the two users.

There are also some challenging technical issues associated with the promising avenues of this work, and we briefly summarize two of them in the following. First, as a result of the curse of dimension, we are limited to considering only a special case where the maximum retransmission is limited to 2, and the power allocation policy is assumed to be discrete in order to solve the MDP problem. However, with our derived BLERs in hand, we could look forward to a more flexible and intelligent transmission schemes which enable unlimited retransmissions and continuous power allocations. Solving such problems by the MDP method requires extremely high computing capacity and large memory, but it could be solved by leveraging the expandability of deep reinforcement learning. Second, to analyze the effect of unreliable and delayed feedback on AoI may also be an interesting work. Third, we notice that the concurrent work usually consider the optimization of NOMA systems as a two-step problem. That is, to first conduct user pairing optimization; and then, within the pairing users, to carry out optimized resource scheduling. In this regard, to jointly optimize the AoI of NOMA systems, wherein the user pairing and resource scheduling can be simultaneously optimized, would be a valuable work.

Appendix A
Proof of Lemma 1

The Laplace transformation of (15) can be approximated by adopting Gaussian-Chebyshev quadrature, given as

$$\begin{aligned}\mathcal{L}_{\gamma_{i2}^{(k)}}(s) &= \int_0^{\theta_k} f_{\gamma_{i2}^{(k)}}(y_k)\, e^{-s y_k} dy_k \\ &\approx c_k^{(i)} \sum_{n_k=1}^{N} \mathcal{G}(a_{n_k})\, e^{\frac{-s\left(\theta_k\left(a_{n_k}+1\right)\right)}{2}}.\end{aligned} \tag{49}$$

Based on (16), we obtain that

$$\begin{aligned}&\mathcal{L}_{\gamma_{i2}(\tilde{t_2})}(s) \\ &= \int_0^{\infty} f_{\gamma_{i2}(\tilde{t_2})}(y)\, e^{-sy} dy \\ &= \prod_{k=1}^{\tilde{t_2}} \int_0^{\infty} f_{\gamma_{i2}^{(k)}}(y_k)\, e^{-s y_k} dy_k = \prod_{k=1}^{\tilde{t_2}} \mathcal{L}_{\gamma_{i2}^{(k)}}(s) \\ &\approx \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \cdots \sum_{n_{\tilde{t_2}}=1}^{N} \left(\prod_{j=1}^{\tilde{t_2}} c_j^{(i)} \mathcal{G}_j^{(i)}(a_{n_j}) \right) \\ &\quad \times e^{-\frac{s}{2} \sum_{j=1}^{\tilde{t_2}} \theta_j\left(a_{n_j}+1\right)}.\end{aligned} \tag{50}$$

Using the Gaver-Stehfest procedure proposed in [26], the inverse Laplace transformation of (50), i.e., the PDF of $\gamma_{i2}\left(\tilde{t_2}\right)$ can be obtained:

$$f_{\gamma_{i2}\left(\tilde{t_2}\right)}(y) \approx \sum_{n_1=1}^{N}\sum_{n_2=1}^{N}\cdots\sum_{n_{\tilde{t_2}}=1}^{N}\left(\prod_{j=1}^{\tilde{t_2}} c_j^{(i)}\mathcal{G}_j^{(i)}\left(a_{n_j}\right)\right) \cdot\frac{\ln(2)}{t}\sum_{k=1}^{2L}\zeta_k e^{-\frac{k\ln(2)}{2y}\sum_{j=1}^{\tilde{t_2}}\theta_j\left(a_{n_j}+1\right)}. \quad (51)$$

Appendix B
Proof of Theorem 1

Use Riemann integral to approximate (13): $\delta_{j,M}\sqrt{M}\int_{v_{j,M}}^{\mu_{j,M}} F_{\gamma_{ij}}(x)\mathrm{d}x = F_{\gamma_{ij}}(\beta_{j,M})$, thus,

$$\varepsilon_{i2}\left(\tilde{t_2}\right) \approx F_{\gamma_{i2}\left(\tilde{t_2}\right)}(\beta_{i,M}) = \int_0^{\beta_{i,M}} f_{\gamma_{i2}\left(\tilde{t_2}\right)}(y)\mathrm{d}y. \quad (52)$$

We transfer the integral variable y to $(x+1)\beta/2$, so it can be derived that

$$\varepsilon_{i2}\left(\tilde{t_2}\right) \approx \frac{\beta_{i,M}}{2}\int_{-1}^{1} f_{\gamma_{i2}}((x+1)\beta_{i,M}/2)\mathrm{d}x. \quad (53)$$

By using Gaussian-Chebyshev quadrature, the integral in (53) can be approximated as:

$$\bar{\varepsilon}_{i2}\left(\tilde{t_2}\right) \approx \sum_{n_1=1}^{N}\sum_{n_2=1}^{N}\cdots\sum_{n_{\tilde{t_2}}=1}^{N}\left(\prod_{j=1}^{\tilde{t_2}} c_j^{(i)}\mathcal{G}_j^{(i)}\left(a_{n_j}\right)\right) \times\sum_{m=1}^{N}\frac{\pi\ln(2)\sqrt{1-a_m}}{N\beta_{i,M}\left(a_m+1\right)} \sum_{k=1}^{2L}\zeta_k e^{-\frac{k\ln(2)}{\beta_i(a_m+1)}\left(\sum_{j=1}^{\tilde{t_2}}\theta_j\left(a_{n_j}+1\right)\right)}. \quad (54)$$

Appendix C
Proof of Theorem 2

Similarly, we apply Riemann integral to approximate (13) such that

$$\varepsilon_{11}\left(\tilde{t_1}\right) \approx F_{\gamma_{11}\left(\tilde{t_1}\right)}(\beta_{1,M}) = \int_0^{\beta_{1,M}} f_{\gamma_{11}\left(\tilde{t_1}\right)}(y)\mathrm{d}y. \quad (55)$$

Applying (23) in (55) results in the BLER as

$$\begin{aligned}
&\bar{\varepsilon}_{11}\left(\tilde{t_1}\right)\\
&\approx \sum_{j=1}^{\kappa} q_j^{|\mathcal{I}_j|}\sum_{n=1}^{|\mathcal{I}_j|}\frac{(-1)^{|\mathcal{I}_j|-n}\int_0^{\beta_{1,M}} e^{-q_j y}y^{n-1}dy}{(n-1)!}\\
&\times\sum_{\substack{\sum_{i=1}^{\kappa} r_i=|\mathcal{I}_j|-n\\ j\neq i, r_i\geq 0}}\prod_{\substack{m=1\\ m\neq j}}^{\kappa}\binom{|\mathcal{I}_m|+r_m-1}{r_m}\frac{q_m^{|\mathcal{I}_m|}}{(q_m-q_j)^{|\mathcal{I}_m|+r_m}}\\
&=\sum_{j=1}^{\kappa} q_j^{|\mathcal{I}_j|}\sum_{n=1}^{|\mathcal{I}_j|}\frac{(-1)^{|\mathcal{I}_j|-n}\Gamma\left(n, q_j\beta_{j,M}\right)}{q_j^n\left(n-1\right)!}\\
&\times\sum_{\substack{\sum_{i=1}^{\kappa} r_i=|\mathcal{I}_j|-n\\ j\neq i, r_i\geq 0}}\prod_{\substack{m=1\\ m\neq j}}^{\kappa}\binom{|\mathcal{I}_m|+r_m-1}{r_m}\frac{q_m^{|\mathcal{I}_m|}}{(q_m-q_j)^{|\mathcal{I}_m|+r_m}},
\end{aligned} \quad (56)$$

where $\Gamma\left(n,x\right) = \int_0^x t^{n-1}e^{-t}dt$ is the lower incomplete Gamma function.

Appendix D
Proof of Theorem 3

The authors in [33] have given the conditions for the existence of average optimal stationary policy. As such, the proof of existence in this paper is to verify that our model is consistent with Assumptions 3.1, 3.2 and 3.3 in [33].

1) There exist positive constants $M > 0$, $\beta < 1$ and b, and a measurable function, $\zeta\left(s\right) \geq 1$ on *S*, such that the reward function of MDP $c\left(s,a\right) = \omega_1\Delta_1 + \omega_2\Delta_2$, $\left|c\left(s,a\right)\right| \leq M\zeta\left(s\right)$ for all state-action pairs $\left(s,a\right)$ and
$$\sum_{s'\in S}\zeta(s')P(s'\mid s,a) \leq \beta\zeta(s) + b. \quad (57)$$
2) There exist two functions, $v_1, v_2 \in B_\zeta(S)$, and some state, $s_0 \in S$, such that $v_1(s) \leq h_\alpha(s) \leq v_2(s)$ for all $s \in S$ and $\alpha \in (0,1)$, where $h_\alpha(s) = V_\alpha^*(s) - V_\alpha^*\left(s_0\right)$ is relative difference of the function of $V_\alpha^*(s)$ and $B_\zeta(S) := \left\{u : \|u\|_\zeta < \infty\right\}$ denotes Banach space, $\|u\|_\zeta := \sup_{s\in S}\zeta(s)^{-1}\left|u\left(s\right)\right|$ denotes the weighted supremum norm.

For the first condition, when $\zeta\left(s\right) = \omega_1\Delta_1 + \omega_2\Delta_2$ and $b > 1$, according to the restrictions on decision-making based on status, we can derive that

1) When $\alpha_F = u_F = 0$, the BS can only transmit new messages to both users, the proof can be seen in the Appendix A of [26];
2) When $\alpha_F \neq 0, u_F = 1$, the BS can transmit the old message to u_1 or new messages to both users. As shown in (58), bottom of the page, we have $\Delta_1 \geq 2$ because $u_F = 1$. In such a case, it can satisfy that $\eta\left(\pi\right) \leq \beta < 1$.
3) When $\alpha_F \neq 0, u_F = 2$, the proof is the same as that in 2).

$$\eta\left(\pi\right) = \max_{\pi}\left\{\frac{\omega_1(\varepsilon_1\left(\pi\right))_1\Delta_1 + \omega_2(\varepsilon_2\left(\pi\right))_1\Delta_2 + 1 - m}{\omega_1\Delta_1 + \omega_2\Delta_2}, \frac{\omega_1\left((\varepsilon_1\left(\pi\right))_2\left(\Delta_1 - 1\right) + 1\right) + \omega_2(\varepsilon_2\left(\pi\right))_1\Delta_2 + 1 - m}{\omega_1\Delta_1 + \omega_2\Delta_2}\right\} \quad (58)$$

$$\eta\left(\pi\right) = \max_{\pi}\left\{\begin{array}{l}\frac{\omega_1(\varepsilon_1(\pi))_1\Delta_1+\omega_2(\varepsilon_2(\pi))_1\Delta_2+1-m}{\omega_1\Delta_1+\omega_2\Delta_2}, \frac{\omega_1\left((\varepsilon_1(\pi))_2(\Delta_1-1)+1\right)+\omega_2(\varepsilon_2(\pi))_1\Delta_2+1-m}{\omega_1\Delta_1+\omega_2\Delta_2},\\ \frac{\omega_1(\varepsilon_1(\pi))_1\Delta_1+\omega_2\left((\varepsilon_2(\pi))_2(\Delta_2-1)+1\right)+1-m}{\omega_1\Delta_1+\omega_2\Delta_2}, \frac{\omega_1\left((\varepsilon_1(\pi))_2(\Delta_1-1)+1\right)+\omega_2\left((\varepsilon_2(\pi))_2(\Delta_2-1)+1\right)+1-m}{\omega_1\Delta_1+\omega_2\Delta_2}\end{array}\right\} \quad (59)$$

4) When $\alpha_F \neq 0, u_F = 3$, the BS can adopt all actions to transmit the messages to both users.
As shown in (59), bottom of the previous page, because $u_F = 3$, we have $\Delta_1, \Delta_2 \geq 2$, and thus it turns out that $\eta(\pi) \leq \beta < 1$.

According to the above derivation, it can be proved that condition 1 is established under the system considered by this paper. For condition 2, firstly, we set the measurable function $\zeta(s) = \omega_1\Delta_1 + \omega_2\Delta_2$, and according to the condition 1, there exist $1 \leq \kappa < \infty$ that $\sum\limits_{s' \in S} \zeta(s')P(s' \mid s, a) \leq \kappa\zeta(s)$ for all state-action pair (s, a), where $s'$is the possible state from s by taking action a. Because of the characteristic of the AoI evolution, by taking any actions, the maximum AoI increment in each state is 1, so under the Markovian and deterministic (MD) decision rules, we can derive that

$$\sum_{s' \in S} \zeta(s')P(s' \mid s, a) \leq \zeta(s)\ + 1 \leq (1 + 1)\,\zeta(s). \quad (60)$$

Based on the verified Assumption 6.10.2 in [29], we have

$$\begin{aligned} \lambda^M \sum_{s' \in S} \zeta(s')P^M(s' \mid s, a) &\leq \lambda^M\left(\zeta(s)\ + M\right) \\ &\leq \lambda^M\left(1 + M\right)\zeta(s) \quad (61) \end{aligned}$$

Therefore, according to the Proposition 6.10.1 in [34], the optimal value function $V_\alpha^*(s)$ can be expressed as

$$|V_\alpha^*(s)| \overset{\pi^*}{=} \frac{1}{1-\gamma}\left[1 + \lambda\kappa + \cdots + (\lambda\kappa)^{M-1}\right]\zeta(s), \quad (62)$$

where $0 \leq \gamma < 1$ and $\pi^* = (a_1^*, \ldots, a_N^*)$, and $a_j^* \in A^{MD}$. The proof of condition 2 is completed.

References

[1] T. P. C. de Andrade, C. A. Astudillo, L. R. Sekijima, and N. L. S. da Fonseca, "The random access procedure in long term evolution networks for the Internet of Things," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 124–131, Mar. 2017.

[2] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 2731–2735.

[3] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *Proc. 8th Annu. IEEE Commun. Soc. Conf. Sensor, Mesh Ad Hoc Commun. Netw.*, Jun. 2011, pp. 350–358.

[4] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Trans. Inf. Theory*, vol. 65, no. 3, pp. 1807–1827, Mar. 2018.

[5] K. S. A. Krishnan and V. Sharma, "Minimizing age of information in a multihop wireless network," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[6] Z. Ding *et al.*, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 185–191, Feb. 2015.

[7] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE 24th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Sep. 2013, pp. 611–615.

[8] Y. Yu, H. Chen, Y. Li, Z. Ding, and B. Vucetic, "On the performance of non-orthogonal multiple access in short-packet communications," *IEEE Commun. Lett.*, vol. 22, no. 3, pp. 590–593, Mar. 2018.

[9] J. Chen *et al.*, "On the adaptive AoI-aware buffer-aided transmission scheme for NOMA networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2021, pp. 1–6.

[10] Q. Wang, H. Chen, Y. Li, and B. Vucetic, "Minimizing age of information via hybrid NOMA/OMA," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2020, pp. 1753–1758.

[11] A. Maatouk, M. Assaad, and A. Ephremides, "Minimizing the age of information: NOMA or OMA?" in *Proc. IEEE Conf. Comput. Commun. Workshops*, Apr. 2019, pp. 102–108, doi: 10.1109/INFCOMW.2019.8845254.

[12] S. Lin, D. J. Costello, and M. J. Miller, "Automatic-repeat-request error-control schemes," *IEEE Commun. Mag.*, vol. CM-22, no. 12, pp. 5–17, Dec. 1984.

[13] M. Xie, Q. Wang, J. Gong, and X. Ma, "Age and energy analysis for LDPC coded status update with and without ARQ," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10388–10400, Oct. 2020.

[14] E. Tuğçe Ceran, D. Gündüz, and A. György, "Average age of information with hybrid ARQ under a resource constraint," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1900–1913, Mar. 2019.

[15] F. Ghanami, G. A. Hodtani, B. Vucetic, and M. Shirvanimoghaddam, "Performance analysis and optimization of NOMA with HARQ for short packet communications in massive IoT," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4736–4748, Mar. 2021.

[16] D. Marasinghe, N. Rajatheva, and M. Latva-Aho, "Block error performance of NOMA with HARQ-CC in finite blocklength," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2020, pp. 1–6.

[17] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5G systems," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1647–1651, Oct. 2015.

[18] J. Choi, "Power allocation for max-sum rate and max-min rate proportional fairness in NOMA," *IEEE Commun. Lett.*, vol. 20, no. 10, pp. 2055–2058, Oct. 2016.

[19] J. Zhu, J. Wang, Y. Huang, S. He, X. You, and L. Yang, "On optimal power allocation for downlink non-orthogonal multiple access systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2744–2757, Dec. 2017.

[20] Y. Sun, E. Uysal-Biyikoglu, R. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," in *Proc. 35th Annu. IEEE Int. Conf. Comput. Commun.*, Apr. 2016, pp. 1–9.

[21] B. Wang, S. Feng, and J. Yang, "To skip or to switch? Minimizing age of information under link capacity constraint," in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2018, pp. 1–5.

[22] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling policies for minimizing age of information in broadcast wireless networks," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2637–2650, Dec. 2018.

[23] H. Pan and S. C. Liew, "Information update: TDMA or FDMA?" *IEEE Wireless Commun. Lett.*, vol. 9, no. 6, pp. 856–860, Jun. 2020.

[24] B. A. G. R. Sharan, S. Deshmukh, S. R. B. Pillai, and B. Beferull-Lozano, "Energy efficient AoI minimization in opportunistic NOMA/OMA broadcast wireless networks," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 2, pp. 1009–1022, Jun. 2022.

[25] J. F. Grybosi, J. L. Rebelatto, and G. L. Moritz, "Age of information of SIC-aided massive IoT networks with random access," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 662–670, Jan. 2022.

[26] D. Cai, Z. Ding, P. Fan, and Z. Yang, "On the performance of NOMA with hybrid ARQ," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 10033–10038, Oct. 2018.

[27] J. M. Meredith, *Study on Downlink Multiuser Superposition Transmission for LTE*, document RP150496, TSG RAN Meeting 67, Mar. 2015.

[28] Z. Ding, P. Fan, and H. V. Poor, "User pairing in non-orthogonal multiple access downlink transmissions," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–5.

[29] S. M. R. Islam, M. Zeng, O. A. Dobre, and K.-S. Kwak, "Resource allocation for downlink NOMA systems: Key techniques and open issues," *IEEE Wireless Commun.*, vol. 25, no. 2, pp. 40–47, Apr. 2018.

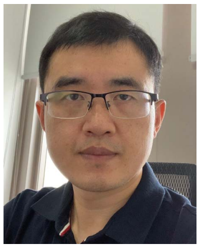
[30] L. Zhu, J. Zhang, Z. Xiao, X. Cao, and D. O. Wu, "Optimal user pairing for downlink non-orthogonal multiple access (NOMA)," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 328–331, Apr. 2019.

[31] B. Makki, T. Svensson, and M. Zorzi, "Finite block-length analysis of the incremental redundancy HARQ," *IEEE Wireless Commun. Lett.*, vol. 3, no. 5, pp. 529–532, Oct. 2014.

[32] E. Levy, "On the density for sums of independent exponential, Erlang and gamma variates," *Stat. Papers*, vol. 63, no. 3, pp. 693–721, Jul. 2021.

[33] X. Guo and Q. Zhu, "Average optimality for Markov decision processes in Borel spaces: A new condition and approach," *J. Appl. Probab.*, vol. 43, pp. 318–334, Jun. 2006.

[34] M. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY, USA: Wiley, 1994.



Shaohua Wu (Member, IEEE) received the Ph.D. degree in communication engineering from the Harbin Institute of Technology in 2009. From 2009 to 2011, he held a post-doctoral position at the Department of Electronics and Information Engineering, Shenzhen Graduate School, Harbin Institute of Technology (Shenzhen), where he has been, since 2012. From 2014 to 2015, he was a Visiting Researcher with BBCR, University of Waterloo, Canada. He is currently a Full Professor with the Harbin Institute of Technology (Shenzhen), China. He is also a Professor with the Peng Cheng Laboratory, Shenzhen, China. He has authored or coauthored over 100 papers in these fields and holds over 40 Chinese patents. His research interests include satellite and space communications, advanced channel coding techniques, space-air-ground-sea integrated networks, and B5G/6G wireless transmission technologies.

Zhihong Deng received the M.S. degree in communication engineering from the Harbin Institute of Technology (Shenzhen), Shenzhen, China, in 2021. His research interests include resource allocation and non-orthogonal multiple access.

Aimin Li (Graduate Student Member, IEEE) received the B.E. degree in electronic and information engineering from the Harbin Institute of Technology (Shenzhen) in 2020, where he is currently pursuing the Ph.D. degree with the Department of Electronic and Information. His research interests include aerospace communications, advanced channel coding techniques, and wireless communications.

Jian Jiao (Member, IEEE) received the M.S. and Ph.D. degrees in communication engineering from the Harbin Institute of Technology (HIT), Harbin, China, in 2007 and 2011, respectively. From 2011 to 2015, he was a Post-Doctoral Research Fellow of 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, The University of Sydney, Sydney, Australia. Since 2015, he has been with the School of Electrical and Information Engineering, HIT, where he is currently an Associate Professor. He is also an Associate Professor with the Peng Cheng Laboratory, Shenzhen. His research interests include error control codes, satellite communications, and massive RA.

Ning Zhang (Senior Member, IEEE) received the B.S. degree from Beijing Jiaotong University, China, in 2007, the M.S. degree from the Beijing University of Posts and Telecommunications in 2010, and the Ph.D. degree from the University of Waterloo, Canada, in 2015. After that, he was a Post-Doctoral Research Fellow at the University of Waterloo and the University of Toronto, Canada, respectively. He is currently an Associate Professor with the University of Windsor, Canada. His research interests include connected vehicles, mobile edge computing, wireless networking, and machine learning. He was a recipient of the Best Paper Awards from the IEEE GLOBECOM in 2014, the IEEE WCSP in 2015, the *Journal of Communications and Information Networks* in 2018, the IEEE ICC in 2019, the IEEE Technical Committee on Transmission Access and Optical Systems in 2019, and the IEEE ICCC in 2019. He serves as an Associate Editor for IEEE Internet of Things Journal, IEEE Transactions on Cognitive Communications and Networking, IEEE Access, and *IET Communications*, and an Area Editor of *Encyclopedia of Wireless Networks* (Springer) and (Cambridge Scholars). He also serves/served as a Guest Editor for several international journals, such as IEEE Wireless Communications, IEEE Transactions on Cognitive Communications and Networking, IEEE Access, and *IET Communications*. He was the Workshop Chair for MobiEdge'18 (in conjunction with IEEE WiMob 2018), CoopEdge'18 (in conjunction with IEEE EDGE 2018), and 5G&NTN'19 (in conjunction with IEEE EDGE 2019).

Qinyu Zhang (Senior Member, IEEE) received the bachelor's degree in communication engineering from the 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 the Shenzhen Graduate School, HIT, 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 the EIE School, HIT. He is also a Professor with the Peng Cheng Laboratory, Shenzhen, China. His research interests include aerospace communications and networks, wireless communications and networks, cognitive radios, signal processing, and biomedical engineering. He has been a TPC Member for the INFOCOM, IEEE ICC, IEEE GLOBECOM, IEEE Wireless Communications and Networking Conference, and other flagship conferences in communications. He was the Associate Chair for Finance of the International Conference on Materials and Manufacturing Technologies in 2012, the TPC Co-Chair of the IEEE/CIC ICCC 2015, the Symposium Co-Chair of the CHINACOM 2011, and the IEEE Vehicular Technology Conference 2016 (Springer). He was the Founding Chair of the IEEE Communications Society Shenzhen Chapter. He is on the Editorial Board of some academic journals, such as *Journal of Communication*, *KSII Transactions on Internet and Information Systems*, and *Science China Information Sciences*.