Age of Information: Can CR-NOMA Help?

Zhiguo Ding©, Fellow, IEEE, Robert Schober©, Fellow, IEEE, and H. Vincent Poor©, Life Fellow, IEEE

Abstract—The aim of this paper is to exploit cognitive-radio inspired NOMA (CR-NOMA) transmission to reduce the age of information in wireless networks. In particular, two CR-NOMA transmission protocols are developed by utilizing the key features of different data generation models and applying CR-NOMA as an add-on to a legacy orthogonal multiple access (OMA) based network. The fact that the implementation of CR-NOMA causes little disruption to the legacy OMA network means that the proposed CR-NOMA protocols can be practically implemented in various communication systems which are based on OMA. Closed-form expressions for the AoI achieved by the proposed NOMA protocols are developed to facilitate performance evaluation, and asymptotic studies are carried out to identify two benefits of using NOMA to reduce the AoI in wireless networks. One is that the use of NOMA provides users more opportunities to transmit, which means that the users can update their base station more frequently. The other is that the use of NOMA can reduce access delay, i.e., the users are scheduled to transmit earlier than in the OMA case, which is useful to improve the freshness of the data available in the wireless network.

Index Terms—Non-orthogonal multiple access (NOMA), age of information (AoI), cognitive-radio NOMA.

I. INTRODUCTION

In order to support the services envisioned for the sixth generation (6G) mobile network, such as ultra massive machine type communications (umMTC) and enhanced ultra-reliable low latency communications (eUuRLLC), it is critical to ensure the freshness of the data collected in the network [1], [2], [3], [4], [5]. For example, as an important application of umMTC, smart cities require the data for air quality control, traffic management, and critical infrastructure monitoring to be timely and frequently collected [6], [7]. We note that the conventional performance evaluation metrics, such as the ergodic data rate and bit error probability, are not adequate for measuring the freshness of the data available in the network, which motivates a recently developed metric, termed the age of information (AoI). In particular, the AoI is defined as the time elapsed between the generation time and the receive time of a successfully delivered update. The AoI achievable for single-user transmission has been rigorously characterized in [1], [2]. For energy constrained wireless networks, such as sensor networks, the use of energy harvesting is important, and the impact of energy harvesting on the AoI has been studied in [8]. For the scenario with correlated information from multiple devices, a new metric, termed correlation-aware AoI, has been developed and optimized for unmanned aerial vehicle (UAV) networks in [9]. Recently, the application of advanced physical layer communication techniques, such as hybrid automatic repeat request (H-ARQ) and cooperative communications, to improve the AoI of wireless networks with one source-destination pair has been considered in [10] and [11], respectively.

For the scenario with multiple users, the AoI analysis is more challenging than that for the single-user scenario. This is due to the fact that multiple users are competing in the same transmission medium, i.e., one user’s update might be preempted by another’s and hence its AoI is affected by the other users’ transmission strategies. The AoI realized by wireless transmission with various random access protocols has been characterized in [12], [13], and [14]. We note that in many wireless networks, the potential collision between multiple users is avoided by applying orthogonal multiple access (OMA) techniques, such as frequency division multiple access (FDMA) and time division multiple access (TDMA). In [15], the impact of these OMA techniques on the AoI was studied, where TDMA was shown to outperform FDMA in terms of the averaged AoI. As non-orthogonal multiple access (NOMA) is more spectrally efficient than OMA, it is natural to consider the use of NOMA for improving the AoI of wireless networks [16]. The authors of [17] focused on a two-user scenario and showed that the spectral efficiency gain of NOMA over OMA indeed can be transferred to a reduction of the AoI. In order to minimize the AoI, a dynamic policy to switch between NOMA and OMA was developed by formulating the AoI minimization problem as a Markov decision process problem. In [18], the impact of stochastic arrivals on the AoI achieved by NOMA was studied, where the performance gain of NOMA over OMA was shown to be significant for large arrival rates. In [19], the AoI of a two-user NOMA...
assisted grant-based network was minimized by applying a Markov decision process framework, and in [20], the AoI of NOMA assisted grant-free transmission was minimized based on an evolutionary game framework. In [21], the AoI realized by reconfigurable intelligent surface (RIS) assisted NOMA transmission was analyzed, and the application of NOMA to reduce the AoI in satellite communications was considered in [22].

This paper considers a general multi-user uplink communication network, where OMA has already been deployed to serve the multiple users, i.e., there is a legacy network based on OMA. Because TDMA outperforms FDMA in terms of AoI, TDMA is considered as an example for OMA in this paper [15]. The aim of the paper is to apply the NOMA principle for AoI reduction without changing the time slot structure of the legacy TDMA network, which is fundamentally different from the existing NOMA schemes in the literature. Take the schemes developed in [23] and [24] as examples, where power-domain NOMA has been used. These existing NOMA schemes require a major modification to the legacy TDMA time frame structure, i.e., the length of the TDMA frame needs to be halved. The property that NOMA is applied without disruptions to legacy networks is valuable since it ensures the compatibility of NOMA with existing multiple access schemes and allows for the possibility to practically implement NOMA without significant modifications of the existing standards. The main contribution of this paper is the application of NOMA to reduce the AoI of OMA based legacy networks, where the adopted forms of NOMA are tailored to the specific features of different data generation models. In addition, analytical results are derived to characterize the AoI achieved by the proposed NOMA protocols by using the average AoI as the metric. Finally, insightful conclusions about the performance gain of NOMA over OMA are obtained, which clearly unveil the benefits of using NOMA for reducing the AoI of wireless networks.

The contributions of the paper can be summarized in detail as follows:

- For the case where each user’s update is generated at the beginning of its transmit time slot, cognitive-radio inspired NOMA (CR-NOMA) is applied to ensure that each user has two opportunities to deliver its update to the base station in each TDMA time frame [25]. A closed-form expression for the AoI realized by CR-NOMA is developed to facilitate the performance evaluation, and an asymptotic analysis reveals that at high signal-to-noise ratio (SNR), CR-NOMA and TDMA realize the same AoI. This conclusion is expected since, at high SNR, each user needs one transmission only to ensure that its update is successfully delivered to the base station. However, at low SNR, the use of CR-NOMA can result in significant AoI reduction compared to TDMA, which demonstrates that one benefit of using NOMA is to offer users more chances to transmit, i.e., by using NOMA the users can update their base station more frequently.

- For the case where each user’s update is generated at the beginning of a TDMA time frame, a modified CR-NOMA protocol is developed to demonstrate another benefit of using NOMA to improve the AoI. In particular, the modified CR-NOMA protocol effectively reduces the users’ access delay, i.e., the users are scheduled to transmit earlier than in the TDMA case, which is useful for improving the freshness of the data available at the base station. For example, a user which is scheduled in a time slot close to the end of a TDMA frame experiences severe access delay and hence suffers from a large AoI with TDMA, because it has to wait for a long time before it can send its update which has been generated at the beginning of the frame. The use of NOMA ensures that this user can jump the queue and transmit earlier than with TDMA. An exact expression for the AoI achieved with the modified CR-NOMA protocol is obtained, and an asymptotic analysis demonstrates that the use of NOMA yields a reduction of the AoI at high SNR, compared to TDMA.

II. SYSTEM MODEL

Assume that there exists an OMA based legacy wireless network, where \( M \) users, denoted by \( U_m, 1 \leq m \leq M \), send their updates to the same base station. Because TDMA yields smaller AoI than FDMA [15], TDMA is used as an example of OMA in this paper. Furthermore, assume that there are \( M \) time slots in each TDMA time frame, and \( U_m \) is scheduled to transmit in the \( m \)-th time slot of each frame with transmit power \( P \), as shown in Fig. 1. Denote the duration of each time slot by \( T \) s, and the start of the \( m \)-th time slot in the \( i \)-th frame by \( t_i^m \), \( 1 \leq m \leq M \). Therefore, each user can deliver one update to the base station every \( MT \) s.

The aim of this paper is to apply NOMA for AoI reduction without changing the structure of the legacy TDMA network, which is different from the existing works, such as [23] and [24], which need to change the length of each TDMA frame. As such, NOMA can be used as an add-on to OMA, which avoids changing the fundamental building blocks of the legacy network and hence makes the integration of NOMA more seamless. We note that the use of NOMA requires the implementation of successive interference cancellation (SIC) and synchronization among the users, which means that the performance gain of NOMA over OMA is achieved at the expense of an increase in system complexity.

A. AoI Model

AoI indicates the freshness of the updates successfully delivered to the base station and can be defined as follows. At time \( t \), denote by \( T_m(t) \) the generation time of the freshest

For the case where the number of users is larger than the number of time slots per frame, multiple time frames can be combined to yield a super-frame. For example, assume that there are 32 users and 8 time slots in one frame. In total, 32 time slots are needed to serve the users by using round-robin TDMA. Therefore, 4 frames can be viewed as one super-frame, i.e., there are \( M = 32 \) time slots in the super-frame to serve 32 users.

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Therefore, U update successfully delivered from U Ding et al.: Age of Information: Can CR-NOMA Help? 6453

network is given by and the normalized overall average AoI of the considered

is offered multiple chances to transmit in a short period. As a user is required to repeatedly generate updates if the user

However, GAW potentially results in high system complexity of the data by always transmitting a freshly generated update. The model, GAW has the advantage of improving the freshness

t frame, the update to be sent in this time slot is generated at

transmit time slot of the user [1], [17], [23]. For example, assumes that a new update is generated right before the

which is applicable regardless of whether the time is slotted, as to be shown in Section III.

B. Data Generation Models

When data is generated has significant impact on the AoI, and the following two types of data generation are considered in this paper:

1) The Generate-at-Will (GAW) Model: This model assumes that a new update is generated right before the transmit time slot of the user [1], [17], [23]. For example, if a user decides to transmit in the m-th time slot of the i-th frame, the update to be sent in this time slot is generated at t_i^m, as shown in Fig. 2(a). As a widely used data generation model, GAW has the advantage of improving the freshness of the data by always transmitting a freshly generated update. However, GAW potentially results in high system complexity as a user is required to repeatedly generate updates if the user is offered multiple chances to transmit in a short period.

2) The Generate-at-Request (GAR) Model: With this model, the base station requests each user to generate an update at the beginning of each time frame, instead of each time slot, as shown in Fig. 2(b). If retransmission is carried out within one frame, the same update will be sent. GAR is crucial for synchronized sensing, and hence important in many Internet of Things (IoT) applications, such as structural health monitoring and autonomous driving [26], [27]. GAR can lead to larger AoI than GAW, since a user’s access delay, i.e., the duration between the generation time of an update and the corresponding transmit time, is included in the calculation of the AoI. However, compared to GAW, GAR can reduce system complexity and energy consumption, since GAR avoids asking the users to repeatedly generate updates for retransmission.

C. CR-NOMA Transmission

CR-NOMA can be used as an add-on to TDMA to improve the freshness of the data collected in the network, as explained in the following. To reduce system complexity, the following simple user pairing scheme is considered. In particular, U_m and U_{m’}, where 1 ≤ m ≤ n and m’ = m + \frac{1}{2}, are paired together to share the spectrum, and are allowed to transmit simultaneously in the m-th and the m’-th time slots of each frame.

The aim of CR-NOMA transmission is to increase the likelihood for each user to deliver one update to the base station every MT s, compared to TDMA. Depending on the used data generation model, the application of CR-NOMA transmission is different. In brief, for GAW, U_m and U_{m’} use the m-th and the m’-th time slots of each frame for their first tries, respectively, as shown in Fig. 2(a). If their first tries are not successful, U_m uses the m’-th time slot of the current frame, and U_{m’} uses the m-th time slot of the next frame for retransmission. For GAR, both the users use the m-th time slot for their first tries and the m’-th time slot for their second tries, as shown in Fig. 2(b). The details of CR-NOMA transmission are provided as follows.

1) CR-NOMA Transmission for GAW: In the m-th time slot of the i-th frame, U_m is treated as the primary user and is scheduled to transmit with transmit power P, in the same manner as in TDMA. If this transmission is not successful, U_m is offered another chance to transmit in the m’-th time slot with transmit power \text{PS}. Because the m’-th time slot has been assigned to U_{m’} in the TDMA model, U_{m’} and U_m are treated as the primary and secondary users in the m’-th time slot, respectively, as in conventional underlay cognitive

\[ \Delta_m(t) = t - T_m(t), \]

\[ \Delta = \frac{1}{M} \sum_{m=1}^{M} \lim_{T \to \infty} \int_0^T \frac{T}{\Delta_m} \Delta_m(t) \, dt, \]

\[ \Delta_m(t) = t - T_m(t), \]

\[ \Delta = \frac{1}{M} \sum_{m=1}^{M} \lim_{T \to \infty} \int_0^T \frac{T}{\Delta_m} \Delta_m(t) \, dt, \]

\[ \Delta = \frac{1}{M} \sum_{m=1}^{M} \lim_{T \to \infty} \int_0^T \frac{T}{\Delta_m} \Delta_m(t) \, dt, \]

\[ \Delta = \frac{1}{M} \sum_{m=1}^{M} \lim_{T \to \infty} \int_0^T \frac{T}{\Delta_m} \Delta_m(t) \, dt, \]

\[ \Delta = \frac{1}{M} \sum_{m=1}^{M} \lim_{T \to \infty} \int_0^T \frac{T}{\Delta_m} \Delta_m(t) \, dt, \]
networks [29]. To ensure that the implementation of NOMA is transparent to the primary user, the secondary user’s signal is decoded in the first stage of SIC at the base station and its data rate needs to be capped [25]. For example, in the $m'$-th time slot, $U_m$ is the secondary user and its data rate, denoted by $R_{m,i}^S$, is capped as follows:

$$R_{m,i}^S \leq \log \left( 1 + \frac{P^S|h_{i,m}^m|^2}{P|h_{m,m'}^m|^2 + 1} \right),$$  \hspace{1cm} (3)$$

where the binary logarithm is used, and $U_n$’s channel gain in the $k$-th time slot of the $i$-th frame is denoted by $h_{i,n}^k$. We note that the noise power is assumed to be normalized, and hence $P$ and $P^S$ are the effective transmit SNRs. The users’ channel gains in different time slots are assumed to be independent and identically distributed (i.i.d.) and follow the complex Gaussian distribution with zero mean and unit variance. The assumption of i.i.d. Rayleigh fading is justified for important applications of IoT, such as smart home and intelligent city, where users may not have line-of-sight connections to the base station and suffer similar large-scale path-losses. We emphasize that the SIC decoding order and the user pairing scheme used in this paper do not depend on the users’ channel conditions. For example, because the $m$-th time slot in each frame is allocated to $U_m$ in TDMA, this user’s signal is always decoded after its partner’s signal, regardless of whether this user’s channel is stronger or weaker than its partner’s channel. We note that the performance of NOMA transmission can be further improved by implementing more sophisticated user pairing schemes and hybrid SIC decoding orders, see, e.g., [30].

In the $m'$-th time slot of the $i$-th frame, $U_{m'}$ is scheduled to transmit as the primary user, in the same manner as in TDMA. If this transmission is not successful, the user will use the $m$-th time slot of the next frame to transmit a new update. Because $U_{m'}$ is the secondary user in the $m$-th time slot, its data rate in this time slot needs to be capped similarly as in (3).

2) CR-NOMA Transmission for GAR: Recall that with GAR, the users’ updates are generated at the beginning of each TDMA frame. On the one hand, in order to reduce the access delay, $U_{m'}$ will not wait for the $m'$-th time slot, but use the $m$-th time slot to deliver its update. We note that $U_m$ is treated as the primary user in the $m$-th time slot, which means that $U_{m'}$’s data rate needs to be capped similarly as in (3). Only if $U_{m'}$’s transmission in the $m$-th time slot is not successful, $U_{m'}$ carries out a retransmission in the $m'$-th time slot of the current frame.

On the other hand, $U_m$’s transmission strategy for GAR in the $m$-th time slot is the same as that for GAW. If $U_m$ requires a retransmission, it will be treated as the secondary user in the $m'$-th time slot, and its achievable data rate in this time slot depends on its partner’s transmission strategy, which is different from that for GAW. In particular, if $U_{m'}$’s transmission in the $m$-th time slot is not successful, a retransmission from $U_m$ in the $m'$-th time slot is needed, and hence $U_m$’s data rate in this time slot is capped as in (3). However, if $U_{m'}$’s transmission in the $m'$-th time slot is successful, $U_{m'}$ keeps silent in the $m'$-th time slot, which means that $U_m$’s achievable data rate in this time slot is simply given by:

$$R_{m,i}^S = \log \left( 1 + \frac{P^S|h_{i,m}^m|^2}{P|h_{m,m'}^m|^2 + 1} \right).$$

It is assumed that a user resends the same update for its retransmission. The AoI of the considered CR-NOMA scheme can be further reduced by applying advanced H-ARQ schemes [31]. For example, for chase combining (CC) based H-ARQ, a user can send two identically encoded updates, and the base station can improve the receive SNR by employing maximum ratio combining. Alternatively, incremental redundancy (IR) based H-ARQ can be used, where the user sends two independently encoded updates, which are jointly decoded by the base station. Because of the space limitations, H-ARQ is not considered in this paper, but constitutes an important direction for future research.

The AoI achieved by CR-NOMA transmission for the two data generation models will be analyzed in the following two sections, respectively.

III. AOI OF CR-NOMA TRANSMISSION FOR GAW

In this section, the AoI achieved by CR-NOMA is studied for the GAW model, where TDMA is used as a benchmarking scheme. The benefit of using CR-NOMA to reduce the AoI can be clearly illustrated with the example shown in Fig. 3. Because GAW is used, the instantaneous AoI is reduced to $T$ whenever an update is successfully delivered to the base station. For the illustrated example, $U_m$ fails to deliver its update to the base station in the $m$-th time slot of the $(i+1)$ frame. With TDMA, the user has to wait until the next TDMA frame. However, with NOMA, the user has another chance for retransmission in the current frame, which is helpful to reduce the AoI.

A. AoI Realized for TDMA

The AoI realized for TDMA can be straightforwardly analyzed as follows. Without loss of generality, we focus on $U_m$’s average AoI realized for TDMA, denoted by $\bar{\Delta}_T^m$. Further denote the number of frames between the $(j-1)$-th and the $j$-th successful updates by $x_j$. In Fig. 3(a), an example with $x_j = 2$ is shown. Therefore, finding the average AoI for TDMA is equivalent to find the area of the shaded region in Fig. 3(a), denoted by $Q_j$, which means that $\Delta_T^m$ can be expressed as follows:

$$\bar{\Delta}_T^m = \lim_{j \to \infty} \frac{\sum_{j=1}^j Q_j}{\sum_{j=1}^j x_j \cdot MT} = T + \frac{MT \cdot E\{X^2\}}{2 \cdot E\{X\}}.$$  \hspace{1cm} (5)$$

where $J$ denotes the total number of the successful updates, $Q_j = x_j \cdot MT^2 + \frac{1}{2} x_j^2 M^2 T^2$, $E\{X\} = \lim_{j \to \infty} \sum_{j=1}^j x_j$, and $E\{X^2\} = \lim_{j \to \infty} \sum_{j=1}^j x_j^2$. Since the users’ channel gains $x_j$ are identical, the calculation of the AoI requires essentially the evaluation of the area underneath the instantaneous AoI curve, which is applicable regardless of whether the time is slotted.
are assumed to be i.i.d., $x_j$ follows the geometric distribution, i.e., the probability mass function of $x_j$ is given by $P(X = x_j) = p_{c}^{x_j-1}(1 - p_c)$, where $p_c$ denotes the probability for the event that a user fails to deliver an update in a time slot with $T$ s. By using the assumptions that each update contains $N$ bits and the users’ channel gains are complex Gaussian distributed, $p_{S} = P(T \log(1 + P|h_{m}^{i}|^2) \leq N) = 1 - e^{-\epsilon}$, where $\epsilon = 2\pi - 1$. As a result, the normalized average AoI achieved by TDMA can be obtained as follows:

$$\bar{\Delta}^{(1)} = \bar{\Delta}^{(2)} m = T + \frac{MT}{2} (2e^{\epsilon} - 1),$$

where the first step follows by the fact that the users experience the same AoI in TDMA, and the second step follows by using the mean and the variance of the geometric distribution.

### B. AoI Realized for CR-NOMA

Similar to TDMA, the AoI realized by CR-NOMA is also related to $Q_j$ and $x_j$ as shown in (5); however, the analysis of the statistical properties of $Q_j$ and $x_j$ is much more challenging than in the TDMA case. The reason is that each user has two transmission opportunities in each TDMA frame, which means that the time interval between the two adjacent successful updates is not always a multiple of $MT$. The following lemma provides a closed-form expression for the AoI realized by CR-NOMA.

**Lemma 1:** For the case of the GAW model, the normalized overall average AoI realized by CR-NOMA is given by

$$\Delta^N = T + \Delta(p_0, p_m, p_m'),$$

where $\Delta(x, y, z)$ is defined as follows:

$$\Delta(x, y, z) = \frac{MT}{4} \left[ \frac{2(y + z)^2(1 + x) + yz(1 - x)^2}{(y + z)^2(1 - x)} \right].$$

$p_m = e^{-\frac{\epsilon}{p_S}}$, $p_m' = \left(1 - e^{-\frac{\epsilon}{p_S}}\right) e^{-\frac{\epsilon}{p_S}} \frac{1}{1 + \frac{\epsilon}{p_S}}$, and $p_0 = \left(1 - e^{-\frac{\epsilon}{p_S}}\right) \left(1 - e^{-\frac{\epsilon}{p_S}}\right) \frac{1}{1 + \frac{\epsilon}{p_S}} + \frac{\epsilon}{p_S} (1 + \frac{\epsilon}{p_S})$.

**Proof:** See Appendix A.

Based on the closed-form expression of the AoI given in Lemma 1, an asymptotic analysis can be carried out to obtain an insightful understanding of the impact of NOMA on the AoI. For example, consider the following high SNR scenario, $P_S = P \to \infty$, which means that $p_m \approx 1 - \frac{\epsilon}{P}$. $p_m'$ can be approximated as follows:

$$p_m' \approx \left(1 - e^{-\frac{\epsilon}{p_S}}\right) e^{-\frac{\epsilon}{p_S}} \frac{1}{1 + \frac{\epsilon}{p_S}} \approx \frac{e^2}{P(1 + \epsilon)}.$$

Furthermore, $p_0$ can be approximated at high SNR as follows:

$$p_0 \approx \left(1 - e^{-\frac{\epsilon}{p_S}}\right) \left(1 - e^{-\frac{\epsilon}{p_S}}\right) \frac{1}{1 + \frac{\epsilon}{p_S}} \approx \frac{e^2}{P(1 + \epsilon)}.$$

By using the high-SNR approximations of $p_0, p_m$ and $p_m'$, the AoI achieved by CR-NOMA can be approximated as follows:

$$\Delta^N \approx T + \frac{MT}{4} \frac{2(p_m + p_m')^2 + p_m p_m'}{p_m + p_m'} \approx T + \frac{MT}{4} \frac{(p_m + p_m')^2}{(p_m + p_m')^2} = T + \frac{MT}{2},$$

where step 1 follows by the fact that $p_m \gg p_m'$ at high SNR.

On the other hand, the AoI realized by TDMA can be approximated at high SNR as follows:

$$\Delta^T \approx T + \frac{MT}{2} (2e^{\epsilon} - 1) \approx T + \frac{MT}{2}.$$

Comparing the AoI shown in (10) and (11), the following corollary can be obtained.

**Corollary 1:** For the case of the GAW model, at high SNR, i.e., $P_S \to \infty$, the normalized average AoI achieved by CR-NOMA is same as that of TDMA.

**Remark 1:** The conclusion shown in Corollary 1 is expected since, at high SNR, $U_m$’s first try, i.e., its transmission in the $m$-th time slot, is almost guaranteed to be successful. Therefore, retransmission is not needed at high SNR,
and CR-NOMA is reduced to TDMA, which explains why the two protocols achieve the same AoI at high SNR. However, it is important to point out that the use of CR-NOMA can result in a significant performance gain over TDMA in the low SNR regime, as shown in the simulation section.

**Remark 2:** As an add-on to TDMA, CR-NOMA can be used without changes to the time slot structure of TDMA; however, the implementation of CR-NOMA requires some modifications to the signalling architecture of the legacy TDMA system. For example, because $U_{m}$ and $U_{m'}$ share the same time slot, the two users have to be synchronized for the implementation of the developed NOMA scheme. In addition, the implementation of CR-NOMA requires the base station to have access to the users’ channel state information (CSI). Because the signals from $U_{m}$ and $U_{m'}$ arrive at the base station simultaneously, the two users’ pilot signals need to be orthogonal to each other in order to avoid pilot contamination. Furthermore, after each user’s first transmission, the base station needs to send a one-bit feedback acknowledgement; such that the user can decide whether a retransmission is needed. In addition, when a user acts as the secondary user, the base station also needs to inform the user about the choice of the data rate specified in (3). As a result, some straightforward modifications of the signalling structure of the legacy TDMA system are needed.

**IV. AoI of CR-NOMA Transmission for GAR**

Unlike the GAW model, the GAR model requires the users to generate their updates at the beginning of each frame, which means that the access delay needs to be included in the AoI. The impact of the access delay on the AoI can be illustrated by using Fig. 4, where $U_{m'}$’s AoI experience is considered. With TDMA, $U_{m'}$ has to wait until the $m'$-th time slot of each frame for its transmission. This means that the user’s instantaneous AoI drops to $m'T$, since its update is generated at the beginning of the frame. With CR-NOMA, $U_{m'}$ can also use the $m$-th time slot of each frame, which results in two benefits for reducing the AoI. One is that with two chances to transmit every $MT$ s, the likelihood of a failed update is reduced, which is similar to the GAR case. The second benefit is that the use of NOMA can effectively reduce the access delay, since $U_{m'}$ can transmit earlier than for TDMA by using the $m$-th time slot instead of the $m'$-th time slot, and its instantaneous AoI can drop to $mT$. As will be shown in the simulation section, this reduction of the access delay is important for AoI reduction particularly at high SNR. Also, this benefit is not affected by possible correlations of the user’s channels.

---

6For GAW, a user acts as a secondary user for its second transmission. For GAR, $U_{m}$ and $U_{m'}$ act as secondary users in the $m'$-th and $m$-th time slots, respectively. We note that for GAR, the first transmission of $U_{m'}$ happens in the $m$-th time slot. If this first try is successful, $U_{m'}$ does not carry out retransmission in the $m$-th time slot, which means that $U_{m}$’s data rate is $\log(1 + P|h_{m}^{i+1}|^2)$ instead of the one shown in (3). We further note that a secondary user can remain silent if the expected data rate is not sufficient to deliver its update, which reduces energy consumption but prevents possible extensions incorporating H-ARQ.
By using steps similar to those in Section III-A, \( U_{m'} \)'s averaged AoI achieved by TDMA for the GAR model can be obtained as follows:

\[
\Delta^\tau_{m'} = m'T + \frac{MT}{2} (2\epsilon - 1).
\]  

(13)

Comparing (5) to (13), one can find that with the GAR model, the users experience a larger AoI, which is mainly due to the access delay, i.e., a user generates its update at the beginning of one frame but has to wait for its turn to deliver the update.

B. AoI Realized for CR-NOMA

The AoI realized by CR-NOMA in the GAR case is challenging to analyze due to the fact that a user’s instantaneous AoI does not always drop to the same value, which is different from the GAW case. In particular, in the GAW case, a user’s instantaneous AoI is always reduced to \( T \), regardless of which time slot is used for update delivery. However, in the GAR case, if \( U_m \) successfully delivers its update in the \( m \)-th time slot of a frame, its instantaneous AoI drops to \( 2T \), whereas its instantaneous AoI drops to \( m'T \) if the transmission in the \( m \)-th time slot is successful. What makes the AoI analysis more complicated is that the calculation of the shaded region \( Q_{m'} \) shown in Fig. 4 does not only depend on which time slot is used for the \((j-1)\)-th successful update but also on which time slot is used for the \( j \)-th successful update. Furthermore, whether a user can successfully deliver its update to the base station is also affected by its partner’s transmission strategy, as discussed in Section II-C. The following lemma provides a closed-form expression for the AoI realized by CR-NOMA in the GAR case.

Lemma 2: For the case of the GAR model, \( U_k \)'s average AoI realized by CR-NOMA is given by

\[
\Delta^N_k = \Delta_{k,0} + \Delta(p_{mk}, p_{nk}, p_{m'k}),
\]

(14)

where \( k \in \{m, m'\}, 1 \leq m \leq M \). \( \Delta_{k,0} \) is given by

\[
\Delta_{k,0} = \frac{(1 - p_{mk})^2}{(p_{mk} + p_{m'k})^2} \times \left\{ \frac{m'Tp_{mk}}{(1 - p_{mk})^2 + \frac{p_{mk}m'Tp_{mk}}{2} - \frac{p_{mk}m'Tp_{mk}}{1 - p_{mk}}} \right\}.
\]

(15)

\[
p_{mm} = e^{-\frac{\tau}{\epsilon}}, \quad p_{mm'} = \left(1 - e^{-\frac{\tau}{\epsilon}} - e^{-\frac{\tau}{\epsilon}}\right) e^{-\frac{\tau}{\epsilon}} + e^{-\frac{\tau}{\epsilon}} \sigma, \quad p_{0mm} = \left(1 - e^{-\frac{\tau}{\epsilon}} - e^{-\frac{\tau}{\epsilon}}\right) \left(1 - e^{-\frac{\tau}{\epsilon}}\right) + e^{-\frac{\tau}{\epsilon}} \sigma (1 - e^{-\frac{\tau}{\epsilon}}), \quad p_{0mm'} = \left(1 - e^{-\frac{\tau}{\epsilon}}\right) e^{-\frac{\tau}{\epsilon}} + e^{-\frac{\tau}{\epsilon}} (1 - e^{-\frac{\tau}{\epsilon}}),
\]

\[
p_{pmm'} = e^{-\frac{\tau}{\epsilon}} \frac{1}{1 + \frac{\tau}{\epsilon}}, \quad p_{m'm'} = \left(1 - e^{-\frac{\tau}{\epsilon}}\right) e^{-\frac{\tau}{\epsilon}} + \frac{\tau}{\epsilon} \frac{1}{P + 1}, \quad \tau = \frac{1 - e^{-\left(\frac{\tau}{\epsilon} + \frac{1}{P + 1}\right)}}{\frac{\tau}{\epsilon} + \frac{1}{P + 1}}.
\]

The normalized overall average AoI realized by CR-NOMA in the GAR case is given by

\[
\bar{\Delta}^N = \frac{1}{M} \sum_{m=1}^{M} \left(\Delta^N_m + \Delta^N_{m'}\right).
\]

Proof: See Appendix B.

Although the AoI expression shown in Lemma 2 is lengthy and complicated, it can be used to obtain an insightful understanding of the impact of NOMA on the AoI. Let us consider the high SNR scenario, i.e., \( P = \infty \). Because the use of NOMA reduces the access delay for \( U_{m'} \), it is expected that \( U_{m'} \)'s AoI experience for TDMA and NOMA will be significantly different, and hence we focus on \( U_{m'} \)'s AoI in the following.

With some straightforward algebraic manipulations, at high SNR, the following approximations can be obtained: \( p_{0mm} = \frac{1}{1 - \frac{\tau}{\epsilon}}, \quad p_{mm'} = \frac{1 - \frac{\tau}{\epsilon}}{1 - \frac{\tau}{\epsilon}}, \quad \text{and} \quad p_{mm'} = \frac{1 - \frac{\tau}{\epsilon}}{1 - \frac{\tau}{\epsilon}} \). On the one hand, by using these approximations, the first term in (14), \( \Delta_{m',0} \), can be approximated at high SNR as follows:

\[
\Delta_{m',0} \approx \frac{1}{(1 - \frac{\tau}{\epsilon})^2} \times \left[ \left(1 - \frac{1 - \frac{\tau}{\epsilon}}{\epsilon} \right) m'T - \frac{\tau}{\epsilon} \right]
\]

\[
+ \frac{\epsilon}{2(1 + \epsilon)} m'T \left(1 - \frac{1 - \frac{\tau}{\epsilon}}{\epsilon} \right)\left(1 - \frac{1 - \frac{\tau}{\epsilon}}{\epsilon} \right) \left(1 - \frac{1 - \frac{\tau}{\epsilon}}{\epsilon} \right) - \frac{1 - \frac{\tau}{\epsilon}}{2(1 + \epsilon)} m'T \left(1 - \frac{1 - \frac{\tau}{\epsilon}}{\epsilon} \right)
\]

\[
= \frac{1}{1 + \epsilon} m'T + \frac{\epsilon}{2} m'T \epsilon(1 + \epsilon) - \frac{1}{2} m'T \epsilon.
\]

(17)

On the other hand, the second term in (14), \( \Delta(p_{0mm}, p_{mm'}, p_{mm'}) \), can be approximated as follows:

\[
\Delta(p_{0mm}, p_{mm'}, p_{mm'}) = \frac{MT}{4}
\]

\[
\frac{2(p_{mm'} + p_{mm'})(1 + p_{0mm})}{p_{mm'} + p_{mm'}(1 - p_{0mm})^2}
\]

\[
\approx \frac{MT}{4} \left(1 - \frac{\tau}{\epsilon}\right)^2 + \frac{1 - \frac{\tau}{\epsilon}}{1 + \frac{\tau}{\epsilon}} \left(1 - \frac{\tau}{\epsilon}\right)
\]

\[
= \frac{MT}{2} \left(1 + \frac{\epsilon}{2(1 + \epsilon)^2}\right).
\]

(18)

Therefore, \( U_{m'} \)'s AoI can be approximated at high SNR as follows:

\[
\bar{\Delta}^N = \Delta_{m',0} + \Delta(p_{0mm'}, p_{mm', p_{mm'}})
\]

\[
\approx \frac{1}{(1 + \epsilon)^2} \left[ m'T(1 + \epsilon) + \frac{\epsilon}{2} m'T + m'T \epsilon(1 + \epsilon) - \frac{1}{2} m'T \epsilon \right]
\]

\[
= \frac{1}{1 + \epsilon} m'T + m'T \epsilon + \frac{1}{2} MT,
\]

(19)

where the last step follows by using the fact that \( m' = m + \frac{M}{2} \).

Recall that \( U_{m'} \)'s AoI for TDMA can be approximated at high SNR as follows: \( m'T + \frac{MT}{2} \). Therefore, the difference between \( U_{m'} \)'s AoI for TDMA and NOMA, denoted by \( D_{m'} \),
where TDMA is used as a benchmarking scheme. The time frame structure shown in Fig. 1 is used, and the users are assumed to have the same target data rate. In addition, the users’ channel gains in different time slots are assumed to be i.i.d. complex Gaussian distributed with zero mean and unit variance for Figs. 5 - 9. Non-i.i.d. channel gains are considered in Fig. 10. For illustration, define $R \triangleq \frac{N}{T}$ and assume that $\mathcal{P} = \mathcal{P}^S$, where $\mathcal{P}$ is the term transmit SNR in the simulation section because the noise power is assumed to be

\[
\begin{align*}
D_{m'} & \triangleq \frac{1}{(1 + \varepsilon)} [m'T + m'Te] + \frac{1}{2} MT - \left( m'T + \frac{MT}{2} \right) \\
& = \frac{1}{(1 + \varepsilon)} [m'T + m'Te] - m'T \\
& = \frac{m'T + m'Te - m'T - m'Te}{(1 + \varepsilon)} = -\frac{MT}{2(1 + \varepsilon)} < 0, \\
& \approx \frac{1}{(1 + \varepsilon)} [m'T + m'Te] - m'T. \\
\end{align*}
\]

which means that $U_{m'}$ experiences less AoI for NOMA than for TDMA.

In order to find a high-SNR approximation for $\Delta_m^N$, the following approximations can be obtained: $\tau \approx \frac{\epsilon}{p}$, $p_{mm} \approx 1 - \frac{\epsilon}{p}$, $p_{mm'n} \approx \frac{1 - \epsilon}{p} + \frac{\epsilon}{p}$, and $p_{mm'} \approx \frac{\epsilon}{p + \epsilon}$. By applying these approximations and also using steps similar to those used for approximating $\Delta_m^N$, $U_m$’s AoI can be approximated at high SNR as follows:

\[
\Delta_m^N = \Delta_{m,0} + \Delta(p_{mm'0}, p_{mm'), p_{mm'}) \approx m'T + \frac{MT}{2},
\]

which is identical to the case of TDMA. Therefore, the following corollary can be obtained.

Corollary 2: For the case with the GAR model, at high SNR, i.e., $\mathcal{P} = \mathcal{P}^S \rightarrow \infty$, $U_m$’s average AoI for CR-NOMA is strictly smaller than that for TDMA. $U_m$’s average AoIs for TDMA and CR-NOMA are identical at high SNR.

Remark 3: For $U_{m'}$, the performance gain of NOMA over TDMA at high SNR is due to the fact that the use of NOMA reduces the access delay, which is particularly important in the GAR case. Recall that in the GAR case, a user’s update is generated at the beginning of the TDMA frame, which means that a user which is scheduled later in the frame suffers from a larger AoI. The use of NOMA always ensures that $U_{m'}$ does not have to wait until the $m'$-th time slot, but can transmit earlier, i.e., in the $m$-th time slot, which leads to the AoI reduction stated in Corollary 2. We note that the use of NOMA does not improve $U_m$’s access delay, which is the reason why $U_m$’s AoIs for TDMA and NOMA are identical at high SNR.

Remark 4: We further note that the AoI reduction due to the use of CR-NOMA is important to improve user fairness. For example, two user which are scheduled in the first and the last time slots of a frame experience significantly different AoI for TDMA, but the use of CR-NOMA can reduce the difference between the users’ AoI experiences.

V. NUMERICAL STUDIES

In this section, the benefits of NOMA transmission regarding the AoI are studied by using computer simulation results, where TDMA is used as a benchmarking scheme. The time frame structure shown in Fig. 1 is used, and the users are assumed to have the same target data rate. In addition, the users’ channel gains in different time slots are assumed to be i.i.d. complex Gaussian distributed with zero mean and unit variance for Figs. 5 - 9. Non-i.i.d. channel gains are considered in Fig. 10. For illustration, define $R \triangleq \frac{N}{T}$ and assume that $\mathcal{P} = \mathcal{P}^S$, where $\mathcal{P}$ is the term transmit SNR in the simulation section because the noise power is assumed to be

\[
\begin{align*}
\Delta_m^N & \approx \Delta_m + \Delta(p_{mm}), p_{mm}) \approx m'T + \frac{MT}{2}, \\
\end{align*}
\]

normalized. Because the AoI depends on the data generation models, the benefits of NOMA for AoI reduction are studied in two different subsections in the following.

A. The Generate-at-Will Model

Recall that for the GAW model, each user generates a new update at the beginning of its transmit time slot. The impact of the different transmission protocols on the AoI is investigated in Fig. 5 by assuming that there are $M = 8$ users, i.e., there are 8 time slots in each TDMA frame. As can be seen from the two subfigures in Fig. 5, the AoI achieved by the proposed NOMA transmission protocol can be significantly lower than that for TDMA. For example, for $R = 1$ bits/s/Hz, $T = 1.5$ s, and a transmit SNR of 0 dB, the AoI realized with TDMA is around 28, and the AoI achieved by CR-NOMA is just 20, which means that the use of CR-NOMA reduces the AoI more than a one-quarter reduction compared to TDMA. However, at high SNR, Fig. 5 shows that TDMA and NOMA yield the same AoI, which confirms Corollary 1. We also note that both subfigures in Fig. 5 verify the accuracy of the analytical results presented in Lemma 1 as well as the approximation result developed in (10).
Fig. 6. The impact of the number of users, $M$, on the AoI, for the generate-at-will model with $R = 1.5 \text{ bits/s/Hz}$ and $T = 0.5 \text{ s}$.

Fig. 5 also shows the impact of $R$ on the AoI performance of the considered transmission protocols. In particular, by comparing Fig. 5(a) and Fig. 5(b), one can observe that increasing $R$ increases the AoI for both transmission protocols. Recall that increasing $R$ for a given value of $T$ means that there are more bits contained in each update, which makes transmission failures more likely and hence increases the AoI. We note that the performance gain of NOMA over TDMA becomes larger for larger $R$. For example, for $T = 1.5 \text{ s}$ and a transmit SNR of $0 \text{ dB}$, the performance gain of NOMA over TDMA for $R = 0.5 \text{ bits/s/Hz}$ is $3$, and this performance gain can be increased to almost $8$ for $R = 1 \text{ bits/s/Hz}$. The subfigures in Fig. 5 also show that the AoI realized by the considered protocols increases with $T$, since increasing $T$ for given $R$ means that there are more bits contained in each update.

In Fig. 6, the impact of the number of users, $M$, on the AoI achieved by the considered transmission protocols is studied. As can be seen from the figure, by increasing $M$, the AoI is increased for both transmission protocols. This observation is expected since with more users in the network, each user has to wait for a longer period of time to be served. In addition, one can also observe that the performance gain of the proposed NOMA protocol over TDMA increases as the number of users, $M$, increases. For example, the performance gap between the two protocols is $5$ for $M = 8$, and increases to $25$ for $M = 32$. This observation means that the proposed NOMA protocol is particularly useful for reducing the AoI of networks with massive connectivity, which is a key use case of the 6G system.

B. The Generate-at-Request Model

Recall that for the GAR model, each user generates its update at the beginning of each TDMA frame, instead of each time slot as in the GAW case. In Fig. 7, the impact of the NOMA transmission protocol on the users’ individual AoI is studied for the GAR model. In particular, Fig. 7(a) focuses on $U_m$’s individual AoI achieved by the two considered transmission protocols, $1 \leq m \leq M$. Unlike for the GAW model, different users experience different Aois for the GAR model, i.e., $U_m$’s AoI is larger than that of $U_i$, $m > i$. This observation is expected since $U_m$’s instantaneous AoI can drop to $mT$ at most, whereas $U_i$’s instantaneous AoI can drop to $iT$. Fig. 7(a) also shows that for $U_m$, the performance gain of NOMA over TDMA is similar to that for the GRW case, e.g., the use of NOMA yields a significant performance gain at low SNR but achieves the same AoI as TDMA at high SNR. This performance gain at low SNR is due to the fact that $U_m$ has a second chance for transmission in each frame, whereas for TDMA, $U_m$ has to rely on a single time slot for its updates.

Fig. 7(b) focuses on $U_m$’s individual AoI achieved for the two transmission protocols. As can be seen from the figure, NOMA outperforms TDMA in all SNR regimes. The reason for this superior performance can be explained as follows. Recall that for TDMA, $U_m$ has to rely on the $m$-th time slot only for sending its update to the base station. The use of the proposed NOMA protocol has two advantages for reducing the AoI. One is that the use of NOMA offers the user two chances to transmit in each TDMA time frame. The other is that the proposed NOMA protocol can schedule $U_m$ to transmit earlier, i.e., completing its update in the $m$-th time slot, instead of waiting for the $m$-th time slot as for TDMA. The latter is crucial for NOMA to outperform TDMA in the high SNR regime, as indicated by Corollary 2. Furthermore,
Fig. 8. Comparison of the AoI experienced by $U_m$ and $U_{m'}$, where the generate-at-request model is used, $m = 1$, $m' = 5$, $M = 8$, $R = 1$ bit/s/Hz and $T = 0.5$ s.

We note that both subfigures of Fig. 7 confirm the accuracy of the developed analytical results shown in Lemma 2.

In Fig. 8, the individual AoI experienced by $U_m$ and $U_{m'}$ is compared. As can be seen from the figure, for the TDMA transmission protocol, $U_{m'}$’s AoI is much larger than that of $U_m$, which is due to the fact that $U_{m'}$ has to wait for the $m'$-th time slot to deliver its update and hence experiences large access delays. By using the proposed NOMA protocol, the difference between the two users’ AoI can be reduced significantly. For example, at high SNR, the difference between the two users’ AoI is 2 with TDMA, and can be halved by applying NOMA. Therefore, the use of NOMA can effectively reduce the difference between the users’ AoI and hence improve user fairness, as discussed in Remark 4.

We also note that Fig. 8 demonstrates the accuracy of the high SNR approximations developed in Section IV-B.

In Fig. 9, the normalized overall AoI is used as the metric to study the performance of the proposed NOMA transmission protocol with the GAR model. We note that for the GAW model, the proposed NOMA protocol has the limitation that it can outperform TDMA in the low SNR regime only, as shown in Figs. 5 and 6. Compared to Figs. 5 and 6, Fig. 9 shows that the proposed NOMA protocol can always outperform TDMA and realize a smaller overall AoI in all SNR regimes. The performance gain of NOMA over TDMA is significant for small $R$, and is reduced by increasing $R$, as shown in the two subfigures of Fig. 9. Furthermore, Fig. 9 also demonstrates that the performance gain of NOMA over TDMA increases...
as the number of the users, \( M \), increases, which is consistent with the observations related to Fig. 6.

In Fig. 10, the AoI performance of NOMA transmission is investigated for the case, where \( U_m \)'s channel gains follow a complex Gaussian distribution with zero mean and variance \( \lambda_1 \), and \( U_{m'} \)'s channel gains follow a complex Gaussian distribution with zero mean and variance \( \lambda_2 \). In particular, if \( \lambda_1 \geq \lambda_2 \), \( U_m \) and \( U_{m'} \) can be viewed as the near and far users in conventional NOMA, respectively. Both Figs. 10(a) and 10(b) show that the use of NOMA always reduces the AoI compared to TDMA, regardless of the values of \( \lambda_1 \) and \( \lambda_2 \), which is consistent with the observations from the previous figures. For GAW, Fig. 10(a) reveals that the AoI performance for \( \lambda_1 = 1 \) and \( \lambda_2 = 0.5 \) is similar to that for \( \lambda_1 = 0.5 \) and \( \lambda_2 = 1 \). This means that for GAW, the AoIs of NOMA and TDMA are not sensitive to the change of the users’ roles as near and far users. However, for GAR, the case with \( U_m \) as a near user is beneficial for AoI reduction, as explained in the following. For GAR, \( U_m \)'s access delay is an important contribution to the overall AoI. With TDMA, \( U_m \) has to wait until the \( m \)'-th time slot in each frame, whereas with NOMA, \( U_m \) can transmit earlier by using the \( m \)'-th time slot. However, in the \( m \)'-th time slot, \( U_{m'} \) acts as a secondary user and suffers strong interference from \( U_m \). If \( U_m \)'s role is changed from a far user to a near user, i.e., changing from \( \lambda_2 = 0.5 \) to \( \lambda_2 = 1 \), its channel condition is improved, and hence it is more likely for \( U_{m'} \) to succeed during the \( m \)'-th time slot, instead of waiting for the \( m \)'-th time slot, which can reduce \( U_{m'} \)'s access delay and hence also reduce the overall AoI, as shown in Fig. 10(b).

VI. CONCLUSION

In this paper, NOMA has been used as an add-on to reduce the AoI of a legacy TDMA network. By using the key features of the two considered data generation models, namely GAW and GAR, two CR-NOMA transmission protocols have been developed to reduce the AoI of the network. Closed-form expressions of the AoI achieved by the proposed NOMA protocols have been derived, and an asymptotic analysis has been carried out to show that the use of NOMA can reduce the AoI due to the following two reasons. First, the use of NOMA provides users more chances to transmit, which ensures that the users can update their base station more frequently. Second, the use of NOMA allows the users to transmit earlier than in the TDMA case, and hence, improve the freshness of the data available at the base station. In this paper, it was assumed that the signals of the secondary users are decoded first at the base station before decoding the primary users’ signals. The use of more dynamic SIC decoding orders can potentially lead to a larger AoI reduction, which is an important direction for future research. In addition, in this paper, the users’ CSI was assumed to be perfectly known for the implementation of NOMA, and perfect SIC was also assumed. If the CSI is imperfect, the resulting channel estimation errors can be treated as additional interference sources modeled by a complex Gaussian distribution [32]. Similarly, if the SIC is imperfect, the residual SIC errors can also be modeled as Gaussian interference sources [33]. In general, the presence of these additional Gaussian interference sources makes the performance analysis mathematically more challenging. Thus, interesting areas for future work include the study of the impact of imperfect CSI and SIC on AoI reduction [34]. Another important direction for future research is the investigation of the tradeoff between energy efficiency and AoI reduction, as with NOMA a user may have to transmit twice in one time frame.

APPENDIX A

PROOF FOR LEMMA 1

It is straightforward to verify that all the users experience the same AoI with CR-NOMA for the GAW case, and therefore, \( U_m \)'s AoI is considered in the remainder of this proof. Unlike the case of TDMA, the duration between the beginning and the end of the \( j \)-th successful update is not always a multiple of \( MT \), since each user has two chances to transmit in each frame. For illustration, assume that \( U_m \)'s \((j-1)\)-th successful update finishes in the \( i \)-th frame. Because \( U_m \) has two chances to transmit in each frame, the following two events are defined based on which of the two time slots is used:

\[ E_{jm} = \{ \text{the } (j-1)\text{-th successful update finishes at the end of the } m \text{-th time slot of a frame, e.g., at } t_{i+1} \} \]
\[ E_{jm'} = \{ \text{the } (j-1)\text{-th successful update finishes at the end of the } m' \text{-th time slot of a frame, e.g., at } t_{i'+1} \} \]

(22)

Denote the time interval between the \((j-1)\)-th and the \(j\)-th successful updates by \( y_j \), \( j \geq 1 \), whose value can be obtained considering by the following four cases:

\[ y_j = \begin{cases} x_jMT, & \text{from } t_{i+1}, \text{to } x_jMT + t_{i+1} \\
 x_jMT, & \text{from } t_{i+1}^{m+1}, \text{to } x_jMT + t_{i+1}^{m+1} \\
 x_jMT + \frac{M}{2}T, & \text{from } t_{i+1}^{m+1}, \text{to } x_jMT + \frac{M}{2}T + t_{i+1}^{m+1} \\
 x_jMT - \frac{M}{2}T, & \text{from } t_{i+1}^{m+1}, \text{to } x_jMT - \frac{M}{2}T + t_{i+1}^{m+1} \end{cases} \]

(23)

where \( x_j \) is defined as the number of frames between the \((j-1)\)-th and the \(j\)-th successful updates, \( x_j \in \mathbb{Z} \), and \( \mathbb{Z} \) denotes the integer set.

Eq. (23) shows that \( y_j = x_jMT \) is caused by two different events. One is that, conditioned on \( E_{jm} \), the user fails to update the base station during the first \((x_j-1)\) frames, but successfully sends an update in the \( m \)-th time slot of the \((x_j+i)\)-th frame, where the user’s \((j-1)\)-th successful update is assumed to occur in the \( i \)-th frame without loss of generality. The other is that, conditioned on \( E_{jm'} \), the user fails to update the base station during the first \((x_j-1)\) frames, but successfully sends an update in the \( m' \)-th time slot of the \((x_j+i)\)-th frame. \( y_j = x_jMT + \frac{M}{2}T \) corresponds to the event that, conditioned on \( E_{jm} \), the user fails to update the base station until the \( m' \)-th time slot of the \((x_j+i)\)-th frame. \( y_j = x_jMT - \frac{M}{2}T \) corresponds to the event that, conditioned on \( E_{jm'} \), the user
fails to update the base station until the $m$-th time slot of the $(x_j + 1)$-th frame. Following the definition of $y_j$ in (23), the following four conditional probabilities can be defined: $p_{j1} = \mathbb{P}(y_j = x_jMT|E_{jm})$, $p_{j2} = \mathbb{P}(y_j = x_jMT + \frac{M}{2}T|E_{jm})$, $p_{j3} = \mathbb{P}(y_j = x_jMT|E_{jm'})$, and $p_{j4} = \mathbb{P}(y_j = x_jMT - \frac{M}{2}T|E_{jm'})$, which will be evaluated later.

$U_m$'s average AoI achieved by NOMA can be expressed as follows:

$$\Delta_N^m = \lim_{J \to \infty} \sum_{j=1}^{J} q_j = \lim_{J \to \infty} \sum_{j=1}^{J} T y_j + \frac{1}{2} y_j^2$$

$$= T + \frac{1}{2} \mathbb{E}\{Y^2\},$$

(24)

where $\mathbb{E}\{Y\} = \lim_{J \to \infty} \frac{1}{J} \sum_{j=1}^{J} y_j$ and $\mathbb{E}\{Y^2\} = \lim_{J \to \infty} \frac{1}{J} \sum_{j=1}^{J} y_j^2$. The remainder of the proof is to evaluate $\mathbb{E}\{Y\}$ and $\mathbb{E}\{Y^2\}$.

Define $p_0$ as the probability of the event that $U_m$ fails to deliver an update in both of the two time slots in one frame. Note that, in the two times slots, $U_m$ assumes different roles for transmission, i.e., $U_m$ is the primary user in the $m$-th time slot and the secondary user in the $m'$-th time slot. By using the date rate constraint in (3) and the assumption that the users' channel gains are i.i.d. complex Gaussian distributed, the probability, $p_0$, can be obtained as follows:

$$p_0 = \mathbb{P}\left(\log \left(1 + P|h_{jm}^m|^2\right) \geq \frac{N}{T}\right) = \mathbb{E}\{Y\}.$$
\[ = M^2 T^2 (p_m + p_{m'})^2 \frac{(1 + p_0)}{(1 - p_0)^3} + \frac{M^2}{2} T^2 p_m p_{m'} \frac{1}{1 - p_0}. \]

(31)

where the last step follows from the following infinite sum of series:

\[ \sum_{j=1}^{\infty} j^2 x^j = x \sum_{j=1}^{\infty} \frac{d}{dx} j^2 x^j = x \sum_{j=1}^{\infty} j x (1 + x) \frac{1}{(1 - x)^3}. \]

(32)

Therefore, \( U_m \)'s average AoI can be calculated as follows:

\[ \Delta_N^N = T + \frac{1}{2} \frac{\mathcal{E}\{Y^2\}}{\mathcal{E}\{Y\}} = T + \frac{1}{2} \times \frac{M^2 T^2 (p_m + p_{m'})^2 \frac{(1 + p_0)}{(1 - p_0)^3} + \frac{M^2}{2} T^2 p_m p_{m'} \frac{1}{1 - p_0}}{M T (p_m + p_{m'})^2 \frac{1}{(1 - p_0)^3}}. \]

(33)

By using the fact that all the users experience the same AoI and with some straightforward algebraic manipulations, the lemma is proved.

**APPENDIX B**

**PROOF FOR LEMMA 2**

For the GAR model, different users experience different AoIs in the proof, the common steps for the analysis of the users’ AoIs are provided first and then the specific results for the users’ individual AoIs are presented.

**A. Generic Expression for the AoI, \( \Delta_N^N \), \( k \in \{m, m'\} \)**

To facilitate the analysis of the AoI, the events in (22) are first modified as follows:

\[ E_{jm} = \{ U_k \text{'s } (j-1)\text{-th successful update finishes at the end of the } m\text{-th time slot of a frame, e.g.}, t_i^{m+1}\}, \]

\[ E_{jm'} = \{ U_k \text{'s } (j-1)\text{-th successful update finishes at the end of the } m'\text{-th time slot of a frame, e.g.}, t_i^{m'+1}\}. \]

(34)

where \( k \in \{m, m'\} \).

Denote by \( y_{jk} \) the time interval between \( U_k \text{'s } (j-1)\text{-th and } j\text{-th successful updates, } k \in \{m, m'\} \). Depending on which of the two events, \( E_{jm} \) and \( E_{jm'} \), happens, the value of \( y_{jk} \) will be different. Furthermore, the height of the rectangle in the shaded region shown in Fig. 4 also depends on the two events, \( E_{jm} \) and \( E_{jm'} \). For example, \( U_m \text{'s instantaneous AoI is reset to } mT \text{ if the user's } (j-1)\text{-th successful update finishes in the } m\text{-th time slot of the last frame, i.e., } E_{jm}', \) occurs. If \( E_{jm}' \), occurs, i.e., the user’s \( (j-1)\text{-th successful update finishes in the } m'\text{-th time slot of the last frame, } U_m \text{'s AoI is reset to } m'T, \) \( U_m \text{'s instantaneous AoI is changed similar to that of } U_m \text{'s AoI. Therefore, the average AoI achieved by CR-NOMA can be expressed as follows:} \]

\[ \Delta_k^N = \lim_{J \to \infty} \frac{\sum_{j=1}^{J} Q_j^k}{\sum_{j=1}^{J} y_{jk}} \]

(35)

where \( k \in \{m, m'\} \), \( Q_j^k \) denotes the area of the shaded shape shown in Fig. 4, and \( 1_E \) is an indicator function, i.e., \( 1_E = 1 \) if event \( E \) happens, otherwise \( 1_E = 0 \). By using steps similar to those in the proof of Lemma 1, \( \Delta_k^N \) can be expressed as follows:

\[ \Delta_k^N = \Delta_{k,0} + \frac{1}{2} \mathcal{E}\{Y^2\} \mathcal{E}\{Y\}, \]

(36)

where \( \mathcal{E}\{Y\} = \lim_{J \to \infty} \frac{1}{J} \sum_{j=1}^{J} y_{jk} \), \( \mathcal{E}\{Y^2\} = \lim_{J \to \infty} \frac{1}{J} \sum_{j=1}^{J} y_{jk}^2 \), and \( \Delta_{k,0} \) is defined as follows:

\[ \Delta_{k,0} = \lim_{J \to \infty} \frac{\sum_{j=1}^{J} (1_{E_{jm}} mT + 1_{E_{jm'}} m'T) y_{jk}}{\sum_{j=1}^{J} y_{jk}}. \]

(37)

To facilitate the performance analysis, denote by \( p_{mk} \) the probability of the event that \( U_k \) fails to deliver an update in both of the two time slots of one frame, by \( p_{mk} \) the probability of the event that \( U_k \) successfully delivers its update in the \( m\text{-th time slot of a frame, and by } p_{mk} \) the probability of the event that \( U_k \) fails in the \( m\text{-th time slot but successfully delivers its update in the } m'\text{-th time slot of the same frame, } k \in \{m, m'\}. \)

By using the same steps in Appendix A, it is straightforward to show that the second term in (36) is simply \( \Delta(p_{mk}, p_{mk}, p_{mk}) \). Therefore, in the remainder of the proof, we focus on the evaluation of the first term in (36), \( \Delta_{k,0} \), as well as the probabilities, \( p_{mk}, p_{mk} \) and \( p_{mk} \), as shown in the following sections.

**B. Evaluation of \( \Delta_{k,0}, k \in \{m, m'\} \)**

By using the expectation \( \mathcal{E}\{Y_k\} \), \( \Delta_{k,0} \) in (36) can be expressed as follows:

\[ \Delta_{k,0} = \lim_{J \to \infty} \frac{\sum_{j=1}^{J} 1_{E_{jm}} mT y_{jk} + \sum_{j=1}^{J} 1_{E_{jm'}} m'T y_{jk}}{\sum_{j=1}^{J} y_{jk}} \]

\[ = \frac{mT}{\mathcal{E}\{Y\}} \lim_{J \to \infty} \frac{\sum_{j=1}^{J} 1_{E_{jm}} y_{jk}}{J} + \frac{m'T}{\mathcal{E}\{Y\}} \lim_{J \to \infty} \frac{\sum_{j=1}^{J} 1_{E_{jm'}} y_{jk}}{J}, \]

(38)

Define the following two conditional expectations:

\[ \mathcal{E}\{Y_k | E_{jm}\} = \lim_{J \to \infty} \frac{\sum_{j=1}^{J} 1_{E_{jm}} mT y_{jk}}{\sum_{j=1}^{J} 1_{E_{jm}}} \] and \[ \mathcal{E}\{Y_k | E_{jm'}\} = \lim_{J \to \infty} \frac{\sum_{j=1}^{J} 1_{E_{jm'}} m'T y_{jk}}{\sum_{j=1}^{J} 1_{E_{jm'}}}. \]

which can be used to simplify the expression for \( \Delta_{k,0} \) as follows:

\[ \Delta_{k,0} = \]

\[ \frac{mT \mathcal{E}\{Y\}}{\mathcal{E}\{Y\}} \mathcal{E}\{Y_k | E_{jm}\} + m'T \mathcal{E}\{E_{jm'}\} \mathcal{E}\{Y_k | E_{jm'}\}. \]

(39)
For illustrative purposes, assume that $U_k$’s $(j - 1)$-th successful update happens in the $i$-th frame. Therefore, $E_{jm}$ means that $U_k$’s $(j - 1)$-th successful update happens in the $m$-th time slot of the $i$-th frame, and hence $y_{jk}$ can have the following two forms:

$$y_{jk} = \begin{cases} t_{jk}, & \text{from } t^{n+1}_{jk} \text{ to } t^{n+1}_{jk} + \frac{M}{2} T, \\ t_{jk} + \frac{M}{2} T, & \text{from } t^{n+1}_{jk} \text{ to } t^{n+1}_{jk} + \frac{M}{2} T + t^{n+1}_{jk} \end{cases},$$

(40)

where $t_{jk} = x_{jk} MT$, and $x_{jk}$ denotes the number of frames between $U_k$’s $(j - 1)$-th and $j$-th successful updates. Therefore, $E\{Y_k|E_{jm}\}$ can be obtained as follows:

$$E\{Y_k|E_{jm}\} = \sum_{j \in Z} x_{jk} MT p^k_{j1}$$

$$+ \sum_{j' \in Z} \left( x_{jk} MT + \frac{M}{2} T \right) p^k_{j2},$$

(41)

where $p^k_{j1} = P(y_{jk} = x_{jk} MT | E_{jm}^k)$ and $p^k_{j2} = P(y_{jk} = x_{jk} MT + \frac{M}{2} T | E_{jm}^k)$, for $k \in \{m, m'\}$.

By using the fact that the users’ channel gains in different time slots are i.i.d., the conditional expectation, $E\{Y_k|E_{jm}\}$, can be rewritten as follows:

$$E\{Y_k|E_{jm}\} = \sum_{j=1}^{\infty} j MT p^j_{0k-1} p_{mk}$$

$$+ \sum_{j=1}^{\infty} \left( j MT + \frac{M}{2} T \right) p^j_{0k-1} p_{mk}$$

$$= \frac{MT(p_{mk} + p_{mk}^2)}{(1 - p_{0k})^2} + \frac{1}{2} \left( \frac{MT p_{mk}}{1 - p_{0k}} \right),$$

(42)

where the first step follows from $p^j_{0k-1} = p^j_{0k-1} p_{mk}$ and $p^j_{0k} = p^j_{0k-1} p_{mk}$, and the last step follows by using (29) and (32).

On the other hand, conditioned on $E_{jm'}$, $y_{jk}$ can have the following two forms:

$$y_{jk} = \begin{cases} t_{jk}, & \text{from } t^{n+1}_{jk} \text{ to } t^{n+1}_{jk} + \frac{M}{2} T, \\ t_{jk} - \frac{M}{2} T, & \text{from } t^{n+1}_{jk} \text{ to } t^{n+1}_{jk} - \frac{M}{2} T + t^{n+1}_{jk} \end{cases}.$$

(43)

Based on the above options for $y_{jk}$, the conditional expectation, $E\{Y_k|E_{jm'}\}$, can be obtained as follows:

$$E\{Y_k|E_{jm'}\} = \sum_{j \in Z} x_{jk} MT p^k_{j3}$$

$$+ \sum_{j' \in Z} \left( x_{jk} MT - \frac{M}{2} T \right) p^k_{j4}$$

$$= \sum_{j=1}^{\infty} j MT p^j_{0k-1} p_{mk}$$

$$+ \sum_{j=1}^{\infty} \left( j MT - \frac{M}{2} T \right) p^j_{0k-1} p_{mk}$$

$$= \frac{MT(p_{mk} + p_{mk}^2)}{(1 - p_{0k})^2} - \frac{1}{2} \left( \frac{MT p_{mk}}{1 - p_{0k}} \right),$$

(44)

where $y_{jk} = P(y_{jk} = x_{jk} MT | E_{jm'}^k) = p^j_{0k-1} p_{mk}$, and $p^k_{j4} = P(y_{jk} = x_{jk} MT - \frac{M}{2} T | E_{jm'}^k) = p^j_{0k-1} p_{mk}$.

By using the two conditional expectations, $\Delta_{k,0}$ can be obtained as follows:

$$\Delta_{k,0} = m TP(E_{jm}^k) E\{Y_k|E_{jm}\} + m' TP(E_{jm'}^k) E\{Y_k|E_{jm'}\}$$

$$= m TP_{mk} E\{Y_k|E_{jm}\} + m' TP_{mk} E\{Y_k|E_{jm'}\}$$

$$= \frac{1}{(p_{mk} + p_{mk}^2)(1 - p_{0k})^2} \left( \frac{MT p_{mk}}{2} \right)$$

$$+ \frac{1}{2} \left( \frac{MT p_{mk}}{1 - p_{0k}} \right),$$

(45)

where the second step follows from the fact that $P(E_{jm}^k) = p_{mk}$, $P(E_{jm'}^k) = p_{mk}$, and the last step follows from the fact that $E\{Y_k\} = MT(p_{mk} + p_{mk}^2)\frac{1}{(1 - p_{0k})^2}$. As can be seen from the above expression, the first term of the users’ AoI expression in (36), $\Delta_{k,0}$, can be explicitly written as a function of $p_{0k}$, $p_{mk}$, and $p_{mk}$, which will be evaluated in the following two subsections for the two users, respectively.

C. Evaluation of $p_{0m}$, $p_{m'm}$, and $p_{m'm}$

Recall that $p_{0m}$ is the probability of the event that $U_m$ fails to deliver an update in both of the two time slots of a given frame. Note that in each of the two time slots, $U_m$ is allowed to transmit in different roles. Further note that if the update from $U_{m'}$ in the $m$-th time slot is successful, $U_{m'}$ will remain silent in the $m'$-th time slot, which means that $U_m$ solely occupies this time slot. By using the data rate constraint in (3), the probability, $p_{0}$, can be expressed as shown in (46), bottom of the next page, where the $i$-th frame is used for illustration.

Because the users’ channel gains in different time slots are assumed to be independent, the event $E_1 \triangleq \{ \log \left(1 + P | h_{m}^i |^2 \right) \leq \frac{N}{T}, \log \left(1 + \frac{p^S | h_{m}^i |^2}{P/h_{m}^i |^2 + 1} \right) \leq \frac{N}{T} \}$ is independent from the event $\{ \log \left(1 + \frac{p^S | h_{m'}^i |^2}{P/h_{m'}^i |^2 + 1} \right) \leq \frac{N}{T} \}$.

Therefore, the probability of $E_1$ can be calculated separately as follows:

$$P(E_1) = P\left( P | h_{m}^i |^2 \leq \epsilon, \frac{p^S | h_{m}^i |^2}{P/h_{m}^i |^2 + 1} \leq \epsilon \right) = \int_{0}^{1} \left( 1 - e^{-\frac{p^S}{\epsilon} (P_x + 1)} \right) e^{-\frac{p^S}{\epsilon} P} \frac{1}{P + 1} dx.$$

Similarly define $E_2 \triangleq \{ \log \left(1 + P | h_{m'}^i |^2 \right) \leq \frac{N}{T}, \log \left(1 + \frac{p^S | h_{m'}^i |^2}{P/h_{m'}^i |^2 + 1} \right) \geq \frac{N}{T} \}$. The probability of $E_2$ can be
evaluated as follows:
\[
\mathbb{P}(E_2) = e^{-\frac{e^{-h_m}}{P} \left( 1 - e^{-\left(\frac{P S}{h_m} + 1\right) h_m} \right) \frac{1}{P}} \cdot e^{-\frac{e^{-h_m}}{P} \left( 1 - e^{-\left(\frac{P S}{h_m} + 1\right) h_m} \right) P + 1}.
\]

By substituting (47) and (48) into (46) and with some algebraic manipulations, probability \( p_{0m} \) can be expressed as follows:
\[
p_{0m} = \left( 1 - e^{-\frac{e^{-h_m}}{P} \left( 1 - e^{-\left(\frac{P S}{h_m} + 1\right) h_m} \right) \frac{P S}{h_m} + 1} \right) \times \left( 1 - e^{-\frac{e^{-h_m}}{P} \left( 1 - e^{-\left(\frac{P S}{h_m} + 1\right) h_m} \right) P + 1} \right).
\]

Recall that \( p_{m'm} \) is the probability for the event that the user fails to deliver its update in the \( m \)-th time slot but successfully delivers the update in the \( m' \)-th time slot. This probability can be evaluated as in (49), shown at the bottom of the page, where the last step is obtained by following steps similar to those for evaluating \( p_{0m} \). It is straightforward to show that \( p_{nm} = p_{m} \).

By substituting the expressions of \( p_{0m}, p_{nm}, p_{m'm} \) and \( \Delta_{m,0} \) in (36), a closed-form expression for \( U_m \)'s AoI can be obtained as shown in the lemma.

**D. Evaluation of \( p_{0m'} \), \( p_{mm'} \) and \( p_{m'm'} \)**

Recall that \( p_{0m'} \) is the probability of the event that \( U_{m'} \) fails to deliver an update in one frame. Again, we take the \( i \)-th frame as an example. Note that \( U_{m'} \) is the secondary user in the \( m \)-th time slot and the primary user in the \( m' \)-th time slot, which means that \( p_{0m'} \) can be expressed as follows:
\[
p_{0m'} = \mathbb{P}( \log \left( 1 + \frac{P S h_{m'}^i}{P |h_m^i|^2 + 1} \right) \leq \frac{N}{T}, \log \left( 1 + \frac{P S h_{m'}^i}{P |h_m^i|^2 + 1} \right) \leq \frac{N}{T}, \log \left( 1 + \frac{P S h_{m'}^i}{P |h_m^i|^2 + 1} \right) \leq \frac{N}{T} )
\]

It is interesting to observe that the expression for \( p_{0m} \) is simpler than that for \( p_{0m} \) in (46) because \( U_{m'} \) is the primary user in the \( m' \)-th time slot and its data rate in this time slot is always \( \log \left( 1 + P |h_m^i|^2 \right) \), regardless of \( U_m \)'s transmission strategy in the \( m' \)-th time slot.

By using the assumption that the users' channels are i.i.d. Rayleigh faded, \( p_{0m'} \) can be evaluated as follows:
\[
p_{0m'} = \mathbb{P}( \log \left( 1 + \frac{P S h_{m'}^i}{P |h_m^i|^2 + 1} \right) \leq \frac{N}{T} ) = e^{-\frac{e^{-h_m}}{P} \left( 1 - e^{-\left(\frac{P S}{h_m} + 1\right) h_m} \right) \frac{1}{P}} \cdot e^{-\frac{e^{-h_m}}{P} \left( 1 - e^{-\left(\frac{P S}{h_m} + 1\right) h_m} \right) P + 1}.
\]

Recall that \( p_{mm'} \) denotes the probability of the event that \( U_{m'} \) successfully delivers its update in the \( m \)-th time slot of a frame, which can be expressed as follows:
\[
p_{mm'} = \mathbb{P}( \log \left( 1 + \frac{P S h_{m'}^i}{P |h_m^i|^2 + 1} \right) \geq \frac{N}{T} ) = e^{-\frac{e^{-h_m}}{P} \left( 1 - e^{-\left(\frac{P S}{h_m} + 1\right) h_m} \right) \frac{1}{P}} \cdot e^{-\frac{e^{-h_m}}{P} \left( 1 - e^{-\left(\frac{P S}{h_m} + 1\right) h_m} \right) P + 1}.
\]

Recall that \( p_{m'm'} \) denotes the probability of the event that \( U_{m'} \) fails in the \( m \)-th time slot but successfully delivers the update in the \( m' \)-th time slot. This probability can be expressed as follows:
\[
p_{m'm'} = \mathbb{P}( \log \left( 1 + \frac{P S h_{m'}^i}{P |h_m^i|^2 + 1} \right) \leq \frac{N}{T} ) = e^{-\frac{e^{-h_m}}{P} \left( 1 - e^{-\left(\frac{P S}{h_m} + 1\right) h_m} \right) \frac{1}{P}} \cdot e^{-\frac{e^{-h_m}}{P} \left( 1 - e^{-\left(\frac{P S}{h_m} + 1\right) h_m} \right) P + 1}.
\]
Robert Schober (Fellow, IEEE) received the Diploma (University) and Ph.D. degrees in electrical engineering from Friedrich-Alexander University Erlangen-Nürnberg (FAU), Germany, in 1997 and 2000, respectively. From 2002 to 2011, he was a Professor and the Canada Research Chair of the University of British Columbia (UBC), Vancouver, Canada. Since January 2012, he has been an Alexander von Humboldt Professor and the Chair of Digital Communication at FAU. His research interests include communication theory, wireless and molecular communications, and statistical signal processing. He received several awards for his work, including the 2002 Heinz Maier Leibnitz Award of the German Science Foundation (DFG), the 2004 Innovations Award of the Vodafone Foundation for Research in Mobile Communications, the 2006 UBC Killam Research Prize, the 2007 Wilhelm Friedrich Bessel Research Award of the Alexander von Humboldt Foundation, the 2008 Charles McDowell Award for Excellence in Research from UBC, the 2011 Alexander von Humboldt Professorship, the 2012 NSERC E. W. R. Stacie Fellowship, the 2017 Wireless Communications Recognition Award by the IEEE Wireless Communications Technical Committee, and the 2021 ACM NanoCom Milestone Award for “Fundamental Contributions to the Modeling, Design, and Analysis of Molecular Communication Systems.” Since 2017, he has been listed as a Highly Cited Researcher by the Web of Science. He is a fellow of the Canadian Academy of Engineering, a fellow of the Engineering Institute of Canada, and a member of the German National Academy of Science and Engineering. He served as the Editor-in-Chief of the IEEE TRANSACTIONS ON COMMUNICATIONS from 2012 to 2015. He currently serves as a member of the Editorial Board of the PROCEEDINGS OF THE IEEE and as a VP Publications for the IEEE Communication Society (ComSoc).

H. Vincent Poor (Life Fellow, IEEE) received the Ph.D. degree in EECS from Princeton University in 1977. From 1977 to 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty at Princeton University, where he is currently the Michael Henry Strater University Professor. From 2006 to 2016, he served as the Dean of Princeton’s School of Engineering and Applied Science. He has also held visiting appointments at several other universities, including most recently at Berkeley and Cambridge. His current research interests include information theory, machine learning and network science, and their applications in wireless networks, and energy systems and related fields. Among his publications in these areas is the recent book Machine Learning and Wireless Communications (Cambridge University Press, 2022). He is a member of the National Academy of Engineering and the National Academy of Sciences and a foreign member of the Chinese Academy of Sciences, the Royal Society, and other national and international academies. He received the IEEE Alexander Graham Bell Medal in 2017.