

NOMA-Assisted Grant-Free Transmission: How to Design Pre-Configured SNR Levels?

Zhiguo Ding, Robert Schober, Bayan Sharif, and H. Vincent Poor

Abstract—An effective way to realize non-orthogonal multiple access (NOMA) assisted grant-free transmission is to first create multiple receive signal-to-noise ratio (SNR) levels and then serve multiple grant-free users by employing these SNR levels as bandwidth resources. These SNR levels need to be pre-configured prior to the grant-free transmission and have great impact on the performance of grant-free networks. The aim of this letter is to illustrate different designs for configuring the SNR levels and investigate their impact on the performance of grant-free transmission, where age-of-information is used as the performance metric. The presented analytical and simulation results demonstrate the performance gain achieved by NOMA over orthogonal multiple access, and also reveal the relative merits of the considered designs for pre-configured SNR levels.

Index Terms—Grant-free transmission, non-orthogonal multiple access (NOMA), age of information (AoI).

I. INTRODUCTION

Grant-free transmission is a crucial feature of the sixth-generation (6G) network to support various important services, including ultra-massive machine type communications (umMTC) and enhanced ultra reliable and low latency communications (eURLLC) [1], [2]. Non-orthogonal multiple access (NOMA) has been recognized as a promising enabling technique to support grant-free transmission, and there are different realizations of NOMA-assisted grant-free transmission. For example, cognitive-radio inspired NOMA can be used to ensure that the bandwidth resources, which normally would be solely occupied by grant-based users, are used to admit grant-free users [3]–[5]. Power-domain NOMA has also been shown effective to realize grant-free transmission, where successive interference cancellation (SIC) is carried out dynamically according to the users’ channel conditions [6]–[8].

The aim of this letter is to focus on the application of NOMA assisted random access for grant-free transmission [9]. Unlike the other aforementioned forms of NOMA, NOMA assisted random access ensures that grant-free transmission can be supported without requiring the base station to have access to the users’ channel state information (CSI). The key idea of NOMA assisted random access is to first create multiple receive signal-to-noise ratio (SNR) levels and then serve users by employing these SNR levels as bandwidth resources. These SNR levels need to be pre-configured prior to the grant-free transmission and have great impact on the performance of grant-free networks. In the literature, there exist two SNR-level designs, termed Designs I and II, respectively. Design I is based on a pessimistic approach and is to ensure that a user’s signal can be still decoded successfully, even if all the remaining users choose the SNR level which contributes the most interference, an unlikely scenario in practice [10], [11].

Z. Ding and Bayan Sharif are with Department of Electrical Engineering and Computer Science, Khalifa University, Abu Dhabi, UAE. Z. Ding is also with Department of Electrical and Electronic Engineering, University of Manchester, Manchester, UK. R. Schober is with the Institute for Digital Communications, Friedrich-Alexander-University Erlangen-Nurnberg (FAU), Germany. H. V. Poor is with the Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544, USA.

Design II is based on an optimistic approach, and assumes that each SNR level is to be selected by at most one user [9]. The advantage of Design II over Design I is that the SNR levels of Design II can be chosen much smaller than those of Design I, and hence are more affordable to the users. The advantage of Design I over Design II is that a collision at one SIC stage does not cause all the SIC stages to fail. The aim of this letter is to study the impact of the two SNR-level designs on grant-free transmission, where the age-of-information (AoI) is used as the performance metric [12], [13]. As the AoI achieved by Design I has been analyzed in [11], this letter focuses on Design II, where a closed-form expression for the AoI achieved by NOMA with Design II and its high SNR approximation are obtained. The presented analytical and simulation results reveal the performance gain achieved by NOMA over orthogonal multiple access (OMA). Furthermore, compared to Design I, Design II is shown to be beneficial for reducing the AoI at low SNR, but suffers a performance loss at high SNR.

II. SYSTEM MODEL

Consider a multi-user uplink network, where M users, denoted by U_m , communicate with the same base station in a grant-free manner. In particular, each user generates its update at the beginning of a time frame which consists of N time slots having duration T seconds each. The users compete for channel access to deliver their updates to the base station, where the probability of a transmission attempt, denoted by \mathbb{P}_{TX} , is assumed to be identical for all users.

With OMA, only a single user can be served in each time slot, whereas the benefit of using NOMA is that multiple users can be served simultaneously. Similar to [11], NOMA assisted random access is adopted for the implementation of NOMA. In particular, the base station pre-configures K receive SNR levels, denoted by P_k , where each user randomly chooses one of the K SNR levels for its transmission with equal probability \mathbb{P}_K . If U_1 chooses P_k , U_1 needs to scale its transmit signal by $\frac{P_k}{|h_1^{j,n}|^2}$, where $h_m^{j,n}$ denotes U_m ’s channel gain in the j -th slot of the n -th frame. If the chosen SNR level is not feasible, i.e., $\frac{P_k}{|h_1^{j,n}|^2} > P$, the user simply keeps silent, where P denotes the user’s transmit power budget. Each user is assumed to have access to its own CSI only, and the users’ channels are assumed to be independent and identically complex Gaussian distributed with zero mean and unit variance.

A. Two Designs to Configure the SNR Levels, P_k

Recall that the SNR levels are configured prior to transmission, which means that the SNR levels cannot be related to the users’ instantaneous channel conditions, but are solely determined by the users’ target data rates. In the literature, there exist two SNR-level designs, as explained in the following. For illustrative purposes, assume that $P_1 \geq \dots \geq P_K$, i.e., SIC is carried out by decoding the user using P_i before decoding the one using P_j , $i < j$. For the trivial case where $M < K$, only the M smallest SNR levels are used.

TABLE I
RECEIVE SNR LEVELS REQUIRED BY THE TWO DESIGNS.

	Design I			Design II		
	$R = 1$ $M = 5$	$R = 1$ $M = 10$	$R = 2$ $M = 5$	$R = 1$ $M = 5$	$R = 1$ $M = 10$	$R = 2$ $M = 5$
P_1	85	820	5655	8	8	192
P_2	21	91	471	4	4	48
P_3	5	10	39	2	2	12
P_4	1	1	3	1	1	3

1) *Design I*: In [10], [11], the receive SNR levels are configured as follows:

$$\log \left(1 + \frac{P_k}{1 + (M-1)P_{k+1}} \right) = R, \quad 1 \leq k \leq K-1, \quad (1)$$

and $\log(1 + P_K) = R$, i.e., $P_K = 2^R - 1$ and $P_k = (2^R - 1)(1 + (M-1)P_{k+1})$, where the users are assumed to have the same target data rate, denoted by R . The rationale behind this design is to ensure successful SIC in the worst case, where one user chooses P_k and all the remaining users choose the SNR level which contributes the most interference, i.e., P_{k+1} . This is a pessimistic assumption since not all the remaining users make a transmission attempt, and it is unlikely for all users to choose the same SNR level.

2) *Design II*: Alternatively, the SNR levels can also be configured as follows [9]¹:

$$\log \left(1 + \frac{P_k}{1 + P_{k+1}} \right) = R, \quad 1 \leq k \leq K-1, \quad (2)$$

and $\log(1 + P_K) = R$, i.e., $P_K = 2^R - 1$ and $P_k = (2^R - 1)(1 + P_{k+1})$. Design II has the drawback that one collision in the i -th SIC stage can cause a failure to the earlier stages, i.e., the j -th SIC stage, $j < i$, since there is more than one interference source for P_i . However, Design II offers the benefit that its SNR levels are less demanding than those for Design I, as can be seen from Table I. Recall that a user has to remain silent if its chosen SNR level is not feasible. Because the SNR levels of Design I are large, these SNR levels cannot be fully used by the users, and hence the number of supported users is smaller than that for Design II.

B. AoI of Grant-Free Transmission

For grant-free transmission, the AoI is an important metric to measure how frequently a user can update the base station. In particular, an effective grant-free transmission scheme needs to ensure that the collisions among the users can be effectively reduced, and the base station can be frequently updated, which makes the AoI an ideal metric. Without loss of generality, U_1 is focused on as the tagged user, and its average AoI is defined as follows [12], [13]:

$$\bar{\Delta} = \lim_{W \rightarrow \infty} \frac{1}{W} \int_0^W \Delta(t) dt, \quad (3)$$

where $\Delta(t)$ denotes the time elapsed since the last successfully delivered update. As the AoI achieved by OMA and NOMA with Design I has been analyzed in [11], the AoI achieved by Design II will be focused on in this letter.

¹A more sophisticated design is to introduce an auxiliary parameter, η , and integrate the two designs shown in (1) and (2) together, i.e., $\log \left(1 + \frac{P_k}{1 + \eta P_{k+1}} \right) = R$, where an interesting direction for future research is to optimize η for AoI reduction.

III. AOI PERFORMANCE ANALYSIS

To facilitate the AoI analysis, denote the time interval between the $(n-1)$ -th and the n -th successful updates by Y_n , and denote the time for the n -th successful update to be delivered to the base station by S_n , $n = 1, 2, \dots$. By using the definition of the AoI, $\bar{\Delta}$ can be expressed as follows [11]: $\bar{\Delta} = \frac{\mathcal{E}\{S_{n-1}Y_n\}}{\mathcal{E}\{Y_n\}} + \frac{\mathcal{E}\{Y_n^2\}}{2\mathcal{E}\{Y_n\}}$, where $\mathcal{E}\{\cdot\}$ denotes the expectation operation. With some algebraic manipulations, $\mathcal{E}\{S_{n-1}Y_n\} = \mathcal{E}\{S_n\}\mathcal{E}\{Y_n\} - \mathcal{E}\{S_n^2\} + \mathcal{E}\{S_n\}^2$, $\mathcal{E}\{Y_n\} = \frac{T^N}{1 - s_0^T \mathbf{P}_M^N \mathbf{1}_{M \times 1}}$ and $\mathcal{E}\{Y_n^2\} = N^2 T^2 \frac{1 + s_0^T \mathbf{P}_M^N \mathbf{1}_{M \times 1}}{(1 - s_0^T \mathbf{P}_M^N \mathbf{1}_{M \times 1})^2} + 2\mathcal{E}\{S_n^2\} - 2\mathcal{E}\{S_n\}^2$, $\mathcal{E}\{S_n\} = T \sum_{l=1}^N l \frac{s_0^T \mathbf{P}_M^{l-1} \mathbf{p}}{1 - s_0^T \mathbf{P}_M^N \mathbf{1}_{M \times 1}}$, $\mathcal{E}\{S_n^2\} = T^2 \sum_{l=1}^N l^2 \frac{s_0^T \mathbf{P}_M^{l-1} \mathbf{p}}{1 - s_0^T \mathbf{P}_M^N \mathbf{1}_{M \times 1}}$, $s_0 = [1 \quad \mathbf{0}_{1 \times (M-1)}]^T$, $\mathbf{0}_{m \times n}$ denotes an all-zero $m \times n$ matrix, $\mathbf{1}_{m \times n}$ is an all-one $m \times n$ matrix, $\mathbf{p} = \mathbf{1}_{M \times 1} - \mathbf{P}_M \mathbf{1}_{M \times 1}$, and \mathbf{P}_M is an $M \times M$ matrix to be explained later.

Recall that the considered access competition among the users can be modelled as a Markov chain with $M+1$ states, denoted by s_j , $0 \leq j \leq M$. In particular, state s_j , $0 \leq j \leq M-1$, denotes that j users succeed in updating the base station, and the tagged user is not one of the successful user. s_M denotes that the tagged user succeeds in updating the base station. The state transition probability, denoted by $P_{j,j+i}$, is the probability from s_j to s_{j+i} , $i \geq 0$ and $j+i \leq M-1$. \mathbf{P}_M is an all zero matrix except its element in the $(j+1)$ -th row and $(j+i+1)$ -th column is $P_{j,j+i}$. The calculation of the state transition probability is directly determined by the transmission strategy. The following lemma provides $P_{j,j+i}$ achieved by NOMA with Design II.

Lemma 1. *The state transition probability, $P_{j,j+i}$, achieved by NOMA with Design II is given by*

$$\begin{aligned} P_{j,j+i} = & \sum_{m=i+1}^{M-j} \binom{M-j}{m} \mathbb{P}_{\text{TX}}^m (1 - \mathbb{P}_{\text{TX}})^{M-j-m} \\ & \times \frac{M-j-i}{M-j} m \cdots (m-i+1) \mathbb{P}_K^m \gamma_i \gamma_{m,i} \\ & + \binom{M-j}{i} \mathbb{P}_{\text{TX}}^i (1 - \mathbb{P}_{\text{TX}})^{M-j-i} \frac{M-j-i}{M-j} i! \mathbb{P}_K^i \gamma_i, \end{aligned} \quad (4)$$

for $0 \leq j \leq M-2$ and $1 \leq i \leq \min\{K, M-1-j\}$, and

$$P_{j,j} = 1 - \sum_{i=1}^{\min\{K, M-j\}} \bar{P}_{j,j+i}, \quad (5)$$

for $0 \leq j \leq M-1$, where

$$\begin{aligned} \bar{P}_{j,j+i} = & \sum_{m=i+1}^{M-j} \binom{M-j}{m} \mathbb{P}_{\text{TX}}^m (1 - \mathbb{P}_{\text{TX}})^{M-j-m} \\ & \times m \cdots (m-i+1) \mathbb{P}_K^m \gamma_i \gamma_{m,i} \\ & + \binom{M-j}{i} \mathbb{P}_{\text{TX}}^i (1 - \mathbb{P}_{\text{TX}})^{M-j-i} i! \mathbb{P}_K^i \gamma_i \end{aligned} \quad (6)$$

$$\begin{aligned} \gamma_i = & \sum_{\substack{k_1 + \dots + k_K = i \\ \max\{k_1, \dots, k_K\} = 1}} (1 - \mathbb{P}_{e,1})^{k_1} \cdots (1 - \mathbb{P}_{e,K})^{k_K}, \quad \gamma_{m,i} = \\ & \sum_{k_1 + \dots + k_K = m-i} \frac{(m-i)!}{k_1! \cdots k_K!} \mathbb{P}_{e,1}^{k_1} \cdots \mathbb{P}_{e,K}^{k_K}, \quad \text{and } \mathbb{P}_{e,k} = 1 - e^{-\frac{P_k}{P}}. \end{aligned}$$

Proof. See Appendix A. \square

At high SNR, i.e., $P \rightarrow \infty$, $\mathbb{P}\left(\frac{P_k}{|h_1^{j,n}|^2} > P\right) \rightarrow 0$, i.e., all SNR levels become affordable to the users, and hence the expressions for the state transition probability can be simplified as shown in the following corollary.

Corollary 1. *At high SNR, $P_{j,j+i}$ can be approximated as follows:*

$$P_{j,j+i} \approx \frac{(M-j-1)!}{(M-j-i-1)!} \mathbb{P}_{\text{TX}}^i (1 - \mathbb{P}_{\text{TX}})^{M-j-i} \mathbb{P}_K^i \bar{\gamma}_i, \quad (7)$$

and

$$\bar{P}_{j,j+i} \approx \frac{(M-j)!}{(M-j-i)!} \mathbb{P}_{\text{TX}}^i (1 - \mathbb{P}_{\text{TX}})^{M-j-i} \mathbb{P}_K^i \bar{\gamma}_i, \quad (8)$$

where $\bar{\gamma}_i = \sum_{\substack{k_1+\dots+k_K=i \\ \max\{k_1,\dots,k_K\}=1}} 1$.

Remark 1: The benefit of using NOMA for AoI reduction can be illustrated based on Corollary 1. For the special case of $i = 1$, the high SNR approximation of $P_{j,j+1}$ can be expressed as follows:

$$P_{j,j+1} \approx (M-j-1) \mathbb{P}_{\text{TX}}^i (1 - \mathbb{P}_{\text{TX}})^{M-j-1} \mathbb{P}_K \bar{\gamma}_i. \quad (9)$$

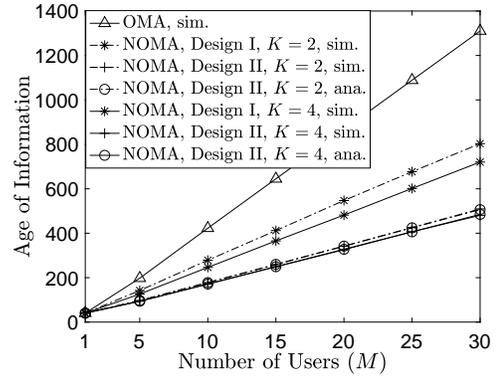
If $\mathbb{P}_K = \frac{1}{K}$, $P_{j,j+1}$ can be simplified as follows:

$$P_{j,j+1} \approx (M-j-1) \mathbb{P}_{\text{TX}}^i (1 - \mathbb{P}_{\text{TX}})^{M-j-1}, \quad (10)$$

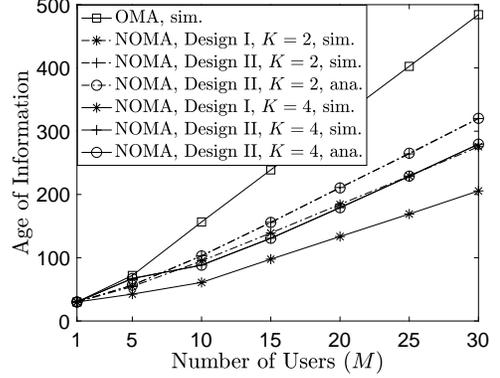
which is exactly the same as that of the OMA case shown in [11]. However, for OMA, $P_{j,j+i} = 0$, $i > 1$, whereas for NOMA, $P_{j,j+i} > 0$, $1 < i \leq K$, which means that with NOMA more users can be served, and hence the AoI of NOMA will be smaller than that of OMA.

IV. SIMULATION RESULTS

In Fig. 1, the AoI of the considered grant-free schemes is shown as a function of the number of users, M . As can be seen from the figure, the use of NOMA transmission can significantly reduce the AoI compared to the OMA case, particularly when there is a large number of users. This ability to support massive connectivity is valuable for umMTC which is the key use case of 6G networks. The figure also demonstrates the accuracy of the analytical results developed in Lemma 1. In addition, Fig. 1(a) shows that at low SNR, the use of Design II yields a significant performance gain over Design I, particularly for large M . However, at high SNR, the use of Design I is more beneficial, as demonstrated in Fig. 1(b). An interesting observation from Fig. 1(b) is that for the special case of $M = 5$ and $P = 30$ dB, the use of $K = 2$ SNR levels yields a better performance than $K = 4$. This is due to the fact that the used choice $\mathbb{P}_{\text{TX}} = \min\left\{\frac{K}{M}, 1\right\}$ is not optimal, as can be explained by using Corollary 1, which shows that $P_{j,j+i}$ is a function of $(1 - \mathbb{P}_{\text{TX}})^{M-1}$ at high SNR. For the special case of $M = 5$ and $K = 4$, $\mathbb{P}_{\text{TX}} = \frac{4}{5}$, and hence $(1 - \mathbb{P}_{\text{TX}})^{M-1}$ can be very small, which causes the AoI of $K = 4$ to be larger than that of $K = 2$. We note that for large M , the performance gain of NOMA over OMA can be always improved by increasing K , i.e., using more SNR levels, as shown in Fig. 1.



(a) $P = 0$ dB



(b) $P = 30$ dB

Fig. 1. AoI achieved by the considered OMA and NOMA schemes as a function of M , where $T = 6$, $R = 1$, $N = 8$, and $K = 2$. For NOMA, $\mathbb{P}_{\text{TX}} = \min\left\{\frac{K}{M}, 1\right\}$ and for OMA, $\mathbb{P}_{\text{TX}} = \frac{1}{M-j}$, where j is the number of users which have successfully delivered their updates to the base station.

In order to better illustrate the impact of the transmit SNR, P , on the performance of the two considered NOMA designs, Fig. 2 shows the AoI as a function of P . Fig. 2 shows that regardless of the choices of the transmit SNR, the AoI of NOMA is always smaller than that of OMA, which is consistent with the observations from Fig. 1. In addition, Fig. 2 also confirms the conclusion that Design II can outperform Design I at low SNR, but suffers a performance loss at high SNR. The reason for Design II to outperform Design I at low SNR is that the SNR levels required by Design I are more demanding than those of Design II, and hence may not be affordable to the users at low SNR, i.e., $\frac{P_k}{|h_1^{j,n}|^2} > P$. The reason for Design I to outperform Design II at high SNR is that, at high SNR, all the levels of the two designs become affordable to the users, and transmission failures are mainly caused by collisions, where unlike Design II, Design I ensures that a collision at the i -th SIC stage does not cause any failure to the j -th stage, $j < i$. An interesting observation from Fig. 2 is that the AoI achieved by Design II may even get degraded by increasing SNR. This is because at low SNR, some users may find that their chosen SNR levels are not affordable, which reduces the number of active users and hence is helpful to reduce the AoI by avoiding collisions.

As discussed previously, the transmission attempt probability, \mathbb{P}_{TX} , is an important parameter for grant-free transmission. In Fig. 3, we show the AoI achieved by the considered schemes for different choices of \mathbb{P}_{TX} . In particular, we consider the

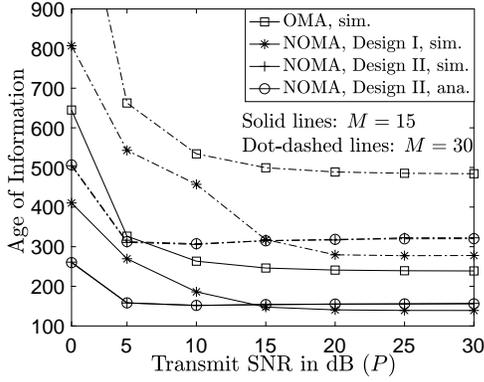
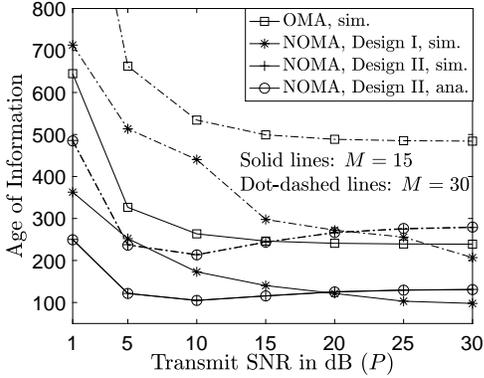
(a) $K = 2$ (b) $K = 4$

Fig. 2. AoI achieved by the considered OMA and NOMA schemes as a function of SNR, P , where $T = 6$, $R = 1$, $N = 8$, and $K = 2$. The same adaptive choices as in Fig. 1 are used for \mathbb{P}_{TX} .

adaptive choice, $\mathbb{P}_{\text{TX}} = \min\{\frac{K}{M}, 1\}$ for NOMA and $\mathbb{P}_{\text{TX}} = \frac{1}{M-j}$ for OMA, where j is the number of users which have successfully delivered their updates to the base station. With the fixed choice, \mathbb{P}_{TX} is set as 0.05. As can be seen from the figure, with a given choice of \mathbb{P}_{TX} , the AoI achieved by NOMA is worse than that of OMA for the special case of low SNR and small M . Nevertheless, the performance gain of NOMA over OMA is still significant in general. In addition, the figure also shows that the use of the adaptive choice of \mathbb{P}_{TX} yields a better performance than that of a fixed \mathbb{P}_{TX} .

V. CONCLUSION

In this letter, the application of NOMA-assisted random access to grant-free transmission has been studied, where the two SNR-level designs and their impact on grant-free networks have been investigated based on the AoI. The presented analytical and simulation results show that the two NOMA designs outperform OMA, and exhibit different behaviours in the low and high SNR regimes.

APPENDIX A PROOF FOR LEMMA 1

Suppose that among the M users, j users have already successfully delivered their updates to the base station, but the tagged user, i.e., U_1 , still has not succeeded. Define $E_{j,j+i}$, $0 \leq i \leq K$, as the event, that i additional users succeed, but the tagged user, U_1 , is not among the i users. The key to studying the AoI is to analyze the state transition probabilities, $\mathbb{P}_{j,j+i} \triangleq \mathbb{P}(E_{j,j+i})$, $0 \leq i \leq K$.

The expressions for $\mathbb{P}_{j,j+i}$, $0 \leq i \leq K$, can be obtained from the following probabilities, $\bar{\mathbb{P}}_{j,j+i} \triangleq \mathbb{P}(\bar{E}_{j,j+i})$, $1 \leq i \leq$

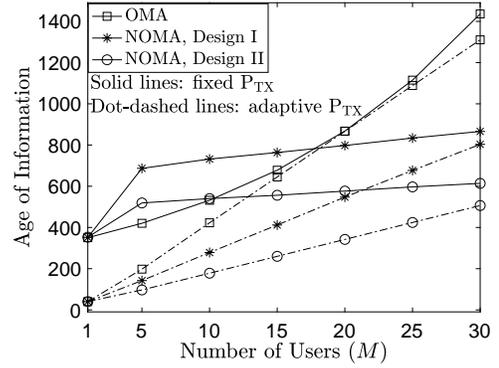
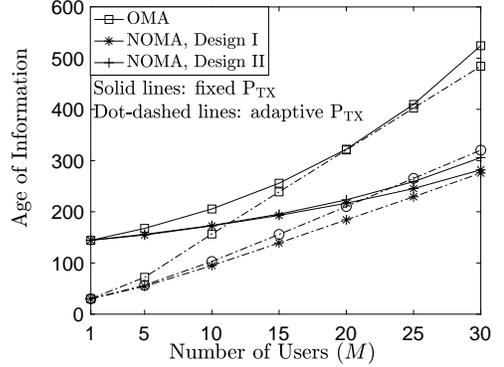
(a) $P = 0$ dB(b) $P = 30$ dB

Fig. 3. AoI achieved by the considered OMA and NOMA schemes with different choices of \mathbb{P}_{TX} , where $T = 6$, $R = 1$, $N = 8$, and $K = 2$.

K , where $\bar{E}_{j,j+i}$ denotes the event, in which among the $M-j$ users, there are i additional users which succeed in updating their base station. Unlike for $E_{j,j+i}$, U_1 can be one of the i successful users for $\bar{E}_{j,j+i}$.

We note that not all the remaining $M-j$ users make a transmission attempt. By using the transmission attempt probability, \mathbb{P}_{TX} , $\bar{\mathbb{P}}_{j,j+i}$ can be expressed as follows:

$$\bar{\mathbb{P}}_{j,j+i} = \sum_{m=i+1}^{M-j} \binom{M-j}{m} \mathbb{P}_{\text{TX}}^m (1 - \mathbb{P}_{\text{TX}})^{M-j-m} \bar{\mathbb{P}}_{m,i}, \quad (11)$$

where $\bar{\mathbb{P}}_{m,i}$ denotes the probability of the event, that among m active users, i.e., m users making a transmission attempt, i users succeed in updating their base station.

Without loss of generality, assume that U_k , $1 \leq k \leq m$, are the m active users. In the following, we focus on a particular event, denoted by E_i , in which among U_k , $1 \leq k \leq m$, U_1, \dots, U_i are the i successful users. Therefore, $\bar{\mathbb{P}}_{m,i}$ can be expressed as follows:

$$\bar{\mathbb{P}}_{m,i} = \binom{m}{i} \mathbb{P}(E_i), \quad (12)$$

where $\binom{m}{i}$ is the number of events which have the same probability as E_i .

If Design I is used, the detection at the n -th SIC stage is affected by the l -th stage, $l < n$, only, and a collision which happens at a later stage, i.e., the p -th stage, $p > n$, has no impact. However, with Design II, a collision will cause all SIC stages to fail, which makes the performance analysis for Design II significantly different from the one shown in [11].

Considering the difference between the two designs, the fact that U_k , $1 \leq k \leq m$, are the m active users, but only U_n , $1 \leq n \leq i$, are successful has the following two implications:

- There is no collision between U_n , $1 \leq n \leq i$, i.e., the i successful users choose different SNR levels. In addition, each user finds its chosen SNR level feasible.
- Each of the failed users, U_n , $i+1 \leq n \leq m$, finds out that its chosen SNR level is not feasible.

The second implication is the key to simplifying the performance analysis, and can be explained as follows. Without loss of generality, assume that U_{i+1} chooses P_k , and finds that P_k is feasible. Because U_{i+1} is one of the active users, it will definitely make an attempt for transmission. Therefore, the only reason to cause this user's transmission to fail is a collision, i.e., another active user chooses the same SNR level as U_{i+1} . This collision at P_k will cause a failure at the k -th SIC stage, as well as the following SIC stages. More importantly, the collision at P_k can also lead to a failure of the early SIC stages, due to the additional interference caused by the two simultaneous transmissions at P_k . Define $E_{i,1}$ as the event where U_n , $1 \leq n \leq i$, successfully deliver their updates to the base station, and $E_{i,2}$ as the event where U_n , $i+1 \leq n \leq m$, fail to deliver their updates to the base station. The two aforementioned implications are also helpful in establishing the independence between the two events, $E_{i,1}$ and $E_{i,2}$, which leads to the following expression:

$$\mathbb{P}(E_i) = \mathbb{P}(E_{i,1})\mathbb{P}(E_{i,2}). \quad (13)$$

In order to better illustrate how $\mathbb{P}(E_i)$ can be evaluated, define $\bar{E}_{i,1}$ as the particular event that U_n chooses P_n , $1 \leq n \leq i$. By using the error probability defined in the lemma, $\mathbb{P}_{e,n}$, $\mathbb{P}(\bar{E}_{i,1})$ can be expressed as follows:

$$\mathbb{P}(\bar{E}_{i,1}) = \mathbb{P}_K^{k_1} (1 - \mathbb{P}_{e,1})^{k_1} \cdots \mathbb{P}_K^{k_K} (1 - \mathbb{P}_{e,K})^{k_K}, \quad (14)$$

where $k_n = 1$, for $1 \leq n \leq i$, and $k_n = 0$ for $i+1 \leq n \leq K$. By using the general expression shown in (14) and enumerating all the possible choices of k_n , $1 \leq n \leq K$, $\mathbb{P}(E_{i,1})$ can be evaluated as follows:

$$\begin{aligned} \mathbb{P}(E_{i,1}) &= i! \sum_{\substack{k_1 + \cdots + k_K = i \\ \max\{k_1, \dots, k_K\} = 1}} \mathbb{P}_K^{k_1} (1 - \mathbb{P}_{e,1})^{k_1} \cdots \mathbb{P}_K^{k_K} (1 - \mathbb{P}_{e,K})^{k_K} \\ &= i! \mathbb{P}_K^i \sum_{\substack{k_1 + \cdots + k_K = i \\ \max\{k_1, \dots, k_K\} = 1}} (1 - \mathbb{P}_{e,1})^{k_1} \cdots (1 - \mathbb{P}_{e,K})^{k_K}, \end{aligned} \quad (15)$$

where $i!$ is the permutation factor since the event where U_1 and U_2 choose P_1 and P_2 , respectively, is different from the event in which U_1 and U_2 choose P_2 and P_1 , respectively.

Similar to $\mathbb{P}(E_{i,1})$, $\mathbb{P}(\bar{E}_{i,2})$ can be obtained as follows:

$$\mathbb{P}(E_{i,2}) = \sum_{\substack{k_1 + \cdots + k_K \\ = m - i}} \frac{(m-i)!}{k_1! \cdots k_K!} \mathbb{P}_K^{k_1} \mathbb{P}_{e,1}^{k_1} \times \cdots \times \mathbb{P}_K^{k_K} \mathbb{P}_{e,K}^{k_K}, \quad (16)$$

where the multinomial coefficients $\frac{(m-i)!}{k_1! \cdots k_K!}$ is needed as explained in the following. Among the $m-i$ unsuccessful users, if k_1 users choose P_1 , there are $\binom{m-i}{k_1}$ possible cases. For the remaining $m-i-k_1$ users, if k_2 users choose P_2 , there are further $\binom{m-i-k_1}{k_2}$ cases. Therefore, the total number of cases for k_n users to choose P_n , $1 \leq n \leq K$, is given by $\binom{m-i}{k_1} \cdots \binom{m-i-k_1-\cdots-k_K}{k_K} = \frac{(m-i)!}{k_1! \cdots k_K!}$. It is interesting to point out that for $E_{i,1}$, the reason for having coefficient $i!$ can

be explained in a similar manner, since $\frac{i!}{k_1! \cdots k_K!} = i!$, if each k_n is either one or zero.

By using (11), (12), and (13), probability $P_{j,j}$ can be expressed as follows:

$$\begin{aligned} P_{j,j} &= 1 - \sum_{i=1}^K \bar{P}_{j,j+i} \\ &= 1 - \sum_{i=1}^K \sum_{m=i+1}^{M-j} \binom{M-j}{m} \mathbb{P}_{\text{TX}}^m (1 - \mathbb{P}_{\text{TX}})^{M-j-m} \\ &\quad \times \binom{m}{i} \mathbb{P}(E_{i,1})\mathbb{P}(E_{i,2}). \end{aligned} \quad (17)$$

By substituting (15) and (16) into (17) and with some algebraic manipulations, the expression for $P_{j,j}$ can be explicitly obtained as shown in the lemma.

By using the difference between $E_{j,j+i}$ and $\bar{E}_{j,j+i}$, probability $P_{j,j+i}$ can be obtained from $\bar{P}_{j,j+i}$ as follows:

$$\begin{aligned} P_{j,j+i} &= \sum_{m=i+1}^{M-j} \binom{M-j}{m} \mathbb{P}_{\text{TX}}^m (1 - \mathbb{P}_{\text{TX}})^{M-j-m} \\ &\quad \times \frac{M-j-i}{M-j} \bar{P}_{m,i}. \end{aligned} \quad (18)$$

The proof of the lemma is complete.

REFERENCES

- [1] X. You, C. Wang, J. Huang *et al.*, "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Sci. China Inf. Sci.*, vol. 64, no. 110301, pp. 1–74, Feb. 2021.
- [2] A. C. Cirik, N. M. Balasubramanya, L. Lampe, G. Vos, and S. Bennett, "Toward the standardization of grant-free operation and the associated NOMA strategies in 3GPP," *IEEE Commun. Standards Mag.*, vol. 3, no. 4, pp. 60–66, Dec. 2019.
- [3] Z. Ding, R. Schober, P. Fan, and H. V. Poor, "Simple semi-grant-free transmission strategies assisted by non-orthogonal multiple access," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4464–4478, Jun. 2019.
- [4] H. Lu, X. Xie, Z. Shi, H. Lei, H. Yang, and J. Cai, "Advanced NOMA assisted semi-grant-free transmission schemes for randomly distributed users," *IEEE Trans. on Wireless Commun.*, to appear in 2023.
- [5] K. Cao, D. Haiyang, B. Wang, L. Lv, J. Tian, Q. Wei, and F. Gong, "Enhancing physical layer security for IoT with non-orthogonal multiple access assisted semi-grant-free transmission," *IEEE Internet of Things Journal*, vol. 9, no. 24, pp. 24 669–24 681, Dec. 2022.
- [6] R. Abbas, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "A novel analytical framework for massive grant-free NOMA," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 2436–2449, Mar. 2019.
- [7] X. Zhang, P. Fan, J. Liu, and L. Hao, "Bayesian learning-based multiuser detection for grant-free NOMA systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 8, pp. 6317–6328, Aug. 2022.
- [8] P. Gao, Z. Liu, P. Xiao, C. H. Foh, and J. Zhang, "Low-complexity block coordinate descend based multiuser detection for uplink grant-free NOMA," *IEEE Trans. Veh. Tech.*, vol. 71, no. 9, pp. 9532–9543, Sept. 2022.
- [9] J. Choi, "NOMA based random access with multichannel ALOHA," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2736–2743, Dec. 2017.
- [10] —, "On throughput bounds of NOMA-ALOHA," *IEEE Wireless Commun. Lett.*, vol. 11, no. 1, pp. 165–168, Jan. 2022.
- [11] Z. Ding, R. Schober, and H. V. Poor, "Impact of NOMA on age of information: A grant-free transmission perspective," *IEEE Trans. Wireless Commun.*, Available on-line at arXiv:2211.13773, 2023.
- [12] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, May 2021.
- [13] H. Zhang, Y. Kang, L. Song, Z. Han, and H. V. Poor, "Age of information minimization for grant-free non-orthogonal massive access using mean-field games," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7806–7820, Nov. 2021.