On Asymmetric Game for NOMA-ALOHA under Fading

Jinho Choi
School of Information Technology
Deakin University
Email: jinho.choi@deakin.edu.au

Youngwook Ko
Department of Electronic Engineering
University of York
Email: youngwook.ko@york.ac.uk

***Abstract*—We consider a non-orthogonal multiple access aided ALOHA (NOMA-ALOHA) that can enhance the spectral efficiency by sending uplink packets of different power levels in wireless multiuser systems. In particular, we develop an asymmetric game theory model for NOMA-ALOHA that involves two different groups of users and analyze the mean rewards and payoffs of actions made by the users. While taking into account not only collisions, but also fading, the derived theoretical results are utilized to formulate a general-sum game with two groups of users and find mixed strategy Nash equilibrium (NE). Interestingly, the analytical and numerical results clearly show that, when the NOMA-ALOHA runs at an NE, a far-user can exploit an improved channel gain of a near-user in the other group in terms of the throughput.**

***Index Terms*—Non-Orthogonal Multiple Access; Random Access; Asymmetric Game; Throughput**

I. Introduction

Non-orthogonal multiple access (NOMA) has been extensively studied to improve the spectral efficiency by exploiting power differences in wireless multiuser systems [1] [2] [3]. For random access systems, in [4], it is also shown that NOMA can help improve the throughput. Since random access does not require coordinated transmissions, it does not have excessive signaling overhead and becomes suitable for machine-type communication (MTC) where a large number of devices of sparse activity are expected to be connected through a shared limited radio resource in various Internet-of-Things (IoT) applications [5] [6].

To understand the performance of random access, game theory is often employed [7] [8] [9]. When NOMA is applied to ALOHA [10] for MTC as in [11] [12], which results in NOMA-ALOHA, the model based on non-cooperative game theory can be used to understand its performance. While non-cooperative game theory is a tool to see the behaviors of players (i.e., devices and sensors that compete for access in random access), it can also be used to derive learning rules for interacting players [13] [14]. In this setup, for each player, say player k, the other players become part of the environment that player k interacts and learns the environment to find the best strategies, possibly through reinforcement learning [15]. For example, in [16], reinforcement learning is studied for ALOHA. Note that although the application of game theory helps understand the performance of random access systems, it is still difficult to find learning rules except for special cases. For example, when all the players have the same conditions (i.e., symmetric games), fictitious play [17], which is a learning rule based on the history of players' selected strategies in the past, can help find best strategies.

In this paper, we consider NOMA-ALOHA game and aim to extend the analysis to more general cases. In particular, after finding the mean rewards under general settings, we focus on the case where players can be divided into two groups. In a cellular system, we can have two groups of users: one group of users are close to the base station (BS) of a cell and the other group of users are far away from the BS. The resulting game becomes an asymmetric game. To characterize its behaviors, we apply the results in [18] and find mixed strategy Nash equilibrium (NE). In particular, we show that a general-sum game can be considered in order to determine the NE of asymmetric games with the two groups of users. Our results show that in the NOMA-ALOHA exploiting an NE, a far-user can exploit an improved channel gain of a near-user in order to increase the throughput.

II. An Overview of NOMA-ALOHA

Suppose that a system consists of multiple users and a BS in a cell. For random access, we assume a time slotted system for slotted ALOHA [10] [19] and a user is to send a packet within a time slot. As in [4], in order to increase the throughput, while a number of power levels can be considered for NOMA, we only consider two power levels, denoted by P_{H} (a high power level) and P_{L} (a low power level), where $P_{\mathrm{H}} > P_{\mathrm{L}} > 0$, in this paper. The resulting random access scheme is referred to as NOMA-ALOHA.

In order to see the throughput improvement of slotted ALOHA by NOMA, suppose that the number of active users follows a Poisson distribution with mean λ. Then, the throughput becomes

$$\eta_{\mathrm{noma}} = \Pr(\text{one active user}) + \Pr(\text{two active users}) \underbrace{\frac{1}{2}}_{(a)} \underbrace{2}_{(b)}$$

This research was supported by the Australian Government through the Australian Research Council's Discovery Projects funding scheme (DP200100391).

$$= \lambda e^{-\lambda} + \frac{\lambda^2}{2!} e^{-\lambda}, \tag{1}$$

where (a) is the probability that one active user chooses $P_{\rm H}$ and the other active user chooses $P_{\rm L}$ and (b) is the number of successfully received packets, which is 2 as one transmits a packet with a transmit power of $P_{\rm H}$ and the other $P_{\rm L}$.

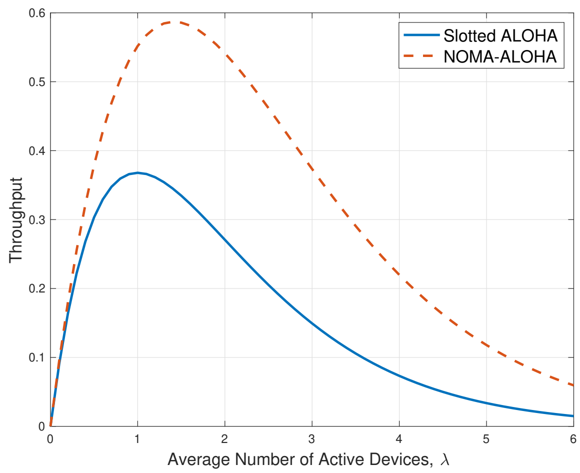


Fig. 1. Throughput curves of slotted ALOHA and NOMA-ALOHA protocols as functions of λ.

Fig. 1 shows the throughput curves of slotted ALOHA and NOMA-ALOHA. Clearly, NOMA-ALOHA performs better than S-ALOHA in terms of throughput. From (1), it can be seen that the throughput of NOMA-ALOHA is maximized when $\lambda = \sqrt{2}$ and the maximum throughput becomes

$$\max \eta_{\rm noma} = (1+\sqrt{2}) e^{-\sqrt{2}} \approx 0.5869.$$

This shows that the maximum throughput of NOMA-ALOHA with two power levels is about 1.6-time higher than that of slotted ALOHA (which is $e^{-1} \approx 0.3679$). In [20], useful bounds on the throughput are shown.

III. Game-Theoretic Analysis

In this section, we consider a game for NOMA-ALOHA and find the average rewards and payoffs under fading.

A. System Model

Suppose that there are K users (or players or agents) and M channels for uplink transmissions. It is assumed that each player has the set of actions, $\mathcal{A} = \{(\mathsf{H},1), \ldots, (\mathsf{H},M), (\mathsf{L},1), \ldots, (\mathsf{L},M), 0\}$, where H and L stand for transmissions with power P_{H} and P_{L}, respectively, and 0 stands for no transmission. Here, $P_{\mathsf{H}} > P_{\mathsf{L}} > 0$ and (H, m) is the action of choosing transmit power P_{H} and channel m. Let $h_{k;m}$ denote the channel coefficient from user k on channel m to the BS. The received signal at the BS on channel m is given by

$$y_m = \sum_{k \in \mathcal{K}_{\mathsf{H},m}} \sqrt{P_{\mathsf{H}}} h_{k;m} s_k + \sum_{k \in \mathcal{K}_{\mathsf{L},m}} \sqrt{P_{\mathsf{L}}} h_{k;m} s_k + n_m, \tag{2}$$

where $\mathcal{K}_{\mathsf{H},m}$ and $\mathcal{K}_{\mathsf{L},m}$ are the index sets of the users who choose channel m with power P_{H} and P_{L}, respectively, s_k is the signal transmitted from user k, and $n_m \sim \mathcal{CN}(0, N_0)$ is the background noise of channel m. Let $\mathbb{E}[s_k] = 0$ and $\mathbb{E}[|s_k|^2] = 1$ for normalization.

We also assume that users do not know the channel coefficients, $h_{k;m}$. As a result, no power control is employed. In addition, each user can choose only one action at a time. Thus, the index sets, $\mathcal{K}_{\mathsf{H},m}$ and $\mathcal{K}_{\mathsf{L},m}$, are disjoint.

B. Formulation of a Game

We can formulate a K-player normal-form game with the following elements:

1) the set of players or users, $\mathcal{K} = \{1, \ldots, K\}$;
2) the set of actions of users, $\mathcal{A}$;
3) the payoffs of players, denoted by R_k, for user k.

To define the payoff, suppose that user k chooses an action of (H, m) or (L, m), which means that this user chooses transmit power P_{H} or P_{L}, respectively, and sends the signal through channel m to the BS. Denote by $V_{k;m}$ and $W_{k;m}$ the instantaneous rewards of user k when choosing (H, m) and (L, m), which become 1 if the transmissions are successful. Otherwise (i.e., transmission is unsuccessful), the instantaneous reward is 0.

Finally, the payoffs can be found as

$$\begin{aligned} R_k(\mathsf{H}, m) &= V_{k;m} - C_{\mathsf{H}} \\ R_k(\mathsf{L}, m) &= W_{k;m} - C_{\mathsf{L}} \\ R_k(0) &= C_0, \end{aligned} \tag{3}$$

where C_i is the cost of action $i \in \{\mathsf{H}, \mathsf{L}, 0\}$. Since the cost depends on the transmit signal power, we expect that $C_{\mathsf{H}} > C_{\mathsf{L}} > C_0$. Note that the payoffs in (3) depend on the others' actions.

For mixed strategies, let

$$\mathbf{x}_k = [x_{k;\mathsf{H},1} \ \ldots \ x_{k;\mathsf{H},M} \ x_{k;\mathsf{L},1} \ \ldots \ x_{k;\mathsf{L},M} \ x_{k;0}]^{\rm T} \in \mathcal{X}, \tag{4}$$

where $x_{k;i,m}$ represent the probability that user k chooses an action of (i, m), $i \in \{\mathsf{H}, \mathsf{L}\}$, and

$$x_{k;0} = 1 - \sum_{m=1}^{M} x_{k;\mathsf{H},m} + x_{k;\mathsf{L},m},$$

which is the probability of no transmission. Here, $\mathcal{X}$ becomes a $2M$-simplex. In addition, let

$$\mathbf{x}_{-k} = (\mathbf{x}_1, \ldots, \mathbf{x}_{k-1}, \mathbf{x}_{k+1}, \ldots, \mathbf{x}_K). \tag{5}$$

C. SINR

Suppose that user k chooses action (i, m), $i \in \{\mathsf{H}, \mathsf{L}\}$. Then, the SINR becomes

$$\mathsf{SINR}_k(i, m) = \frac{\alpha_{k;m} P_i}{I_m}, \ i \in \{\mathsf{H}, \mathsf{L}\}, \tag{6}$$

where $\alpha_{k;m} = |h_{k;m}|^2$ and

$$I_m = \sum_{k' \neq k} \alpha_{k',m} \left(P_{\mathsf{H}} Z_{k';\mathsf{H},m} + P_{\mathsf{L}} Z_{k';\mathsf{L},m} \right) + N_0. \tag{7}$$

Here, $Z_{k;i,m}$, $i \in \{\mathsf{H}, \mathsf{L}\}$ are the activity variables that depend on the action selected by user k and are given by

$$Z_{k;i,m} = \begin{cases} 1, & \text{if user } k \text{ chooses an action of } (i, m) \\ 0, & \text{o.w.} \end{cases} \quad (8)$$

Clearly, $\mathbb{E}[Z_{k;i,m}] = x_{k;i,m}$ and $\sum_{m=1}^M Z_{k;\mathsf{H},m} + Z_{k;\mathsf{L},m} \le 1$.

Throughout the paper, we consider the following assumption.

A1) Independent Rayleigh fading channels are assumed for $|h_{k;m}|$. In particular, we have

$$\alpha_{k;m} \sim \text{Exp}(\bar{\alpha}_{k;m}), \quad (9)$$

where $\bar{\alpha}_{k;m} = \mathbb{E}[\alpha_{k;m}]$

Consequently, we can see that the SINR in (6) is a random variable that depends on the selection of all the users' actions and channel gains.

It is worthy to note that the approaches [4] [11] [21] do not need to consider the SINR in (6), as it is assumed that users know the CSI. If the CSI is known at a user, power control can be performed so that a required SINR for successful decoding can be achieved if no collision happens. However, as in (2), no power control is used. As a result, in order to find the average payoffs, we need to take into account not only collisions, but also fading (i.e., random channel coefficients, (9)).

D. Mean Rewards

In this subsection, we find the mean rewards for given opponents' mixed strategies.

1) Mean Reward with (H, m)*:* Suppose that user k is the player of interest. The signal transmitted by user k can be successfully decoded under the following conditions:

Ea1) user k is only the user choosing (H, m);
Ea2) and the SINR is higher than or equal to Γ_{H}.

For convenience, let

$$\beta_{k;i,m} = \frac{\Gamma_i}{P_i \bar{\alpha}_{k;m}}, \ i \in \{\mathsf{H}, \mathsf{L}\}. \quad (10)$$

We can find the mean reward when user k chooses (H, m) for given $\mathbf{x}_{-k}$ as follows.

Lemma 1: Under the assumption of **A1**, for given $\mathbf{x}_{-k}$, it can be shown that

$$\begin{aligned} \mathbb{E}[V_{k;m}] &= e^{-\beta_{k;\mathsf{H},m} N_0} \prod_{k' \ne k} \phi_{k';m} (1 - x_{k';\mathsf{H},m}) \\ &= e^{-\beta_{k;\mathsf{H},m} N_0} \\ &\times \prod_{k' \ne k} \left(1 - x_{k';\mathsf{H},m} - \frac{\beta_{k;\mathsf{H},m} P_{\mathsf{L}} \bar{\alpha}_{k';m} x_{k';\mathsf{L},m}}{1 + \beta_{k;\mathsf{H},m} P_{\mathsf{L}} \bar{\alpha}_{k';m}} \right), \end{aligned} \quad (11)$$

where

$$\phi_{k';m} = 1 - \frac{\beta_{k;\mathsf{H},m} P_{\mathsf{L}} \bar{\alpha}_{k';m}}{1 + \beta_{k;\mathsf{H},m} P_{\mathsf{L}} \bar{\alpha}_{k';m}} \frac{x_{k';\mathsf{L},m}}{1 - x_{k';\mathsf{H},m}}. \quad (12)$$

Proof: See [22]. ■

2) Mean Reward with (L, m)*:* In this case, the signal transmitted by user k can be successfully decoded under the following conditions:

Eb1) user k is only the user choosing (L, m);
Eb2) at most one another user, say user k', chooses (H, m);
Eb3) and the signals from users k and k' (if exists) can be coded. Γ_{H}.

That is, $W_{k;m} = 1$ if all the above conditions are satisfied. The mean reward can be found as follows.

Lemma 2: Under the assumption of **A1**, for given $\mathbf{x}_{-k}$, the average reward when user k chooses action (L, m) is given by

$$\begin{aligned} \mathbb{E}[W_{k;m}] &= e^{-\beta_{k;\mathsf{L},m} N_0} \prod_{k' \ne k} (1 - x_{k';m}) \\ &\quad \times \left(1 + \sum_{n \ne k} \frac{x_{n;\mathsf{H},m}}{1 - x_{n;m}} \theta_{k,n;m} \right) \\ &= e^{-\beta_{k;\mathsf{L},m} N_0} \left[\prod_{k' \ne k} (1 - x_{k';m}) \right. \\ &\quad \left. + \sum_{n \ne k} \left(\prod_{k' \ne k,n} (1 - x_{k';m}) \right) x_{n;\mathsf{H},m} \theta_{k,n;m} \right], \end{aligned} \quad (13)$$

where $x_{k;m} = x_{k;\mathsf{H},m} + x_{k;\mathsf{L},m}$ and

$$\theta_{k,k';m} = \frac{e^{-\frac{\Gamma_{\mathsf{H}}(\Gamma_{\mathsf{L}}+1)N_0}{P_{\mathsf{H}} \bar{\alpha}_{k';m}}}}{1 + \Gamma_{\mathsf{H}} \frac{P_{\mathsf{L}} \bar{\alpha}_{k;m}}{P_{\mathsf{H}} \alpha_{k';m}}} = \frac{e^{-\beta_{k';\mathsf{H},m}(\Gamma_{\mathsf{L}}+1)N_0}}{1 + \beta_{k';\mathsf{H},m} P_{\mathsf{L}} \bar{\alpha}_{k;m}}.$$

Proof: See [22]. ■

It is noteworthy that we do not consider the capture effect in finding the average rewards as shown above. For coherent decoding, the BS needs to estimate the channel coefficients. To this end, each user can send a pilot signal prior to data packet. In NOMA, there can be two different pilot sequences: one for H (i.e., the case of high transmit power P_{H}) and the other for L (i.e., the case of low transmit power P_{L}). We may assume that the two pilot sequences are orthogonal. The BS can use two correlators with the pilot sequences to estimate the channel coefficients. If there are multiple users that choose H, they will transmit the pilot sequence for H. Then, the channel coefficient as the output of the correlator with the pilot sequence for H becomes a superposition of multiple channel coefficients. Although only one user's channel gain is sufficiently strong (to exploit the capture effect), since the estimated channel coefficient has the other users' channel coefficients, the performance of coherent decoding would be degraded (due to the channel estimation error), which will likely lead to unsuccessful decoding. From this, in above, we only considered the case that there is only one signal (see the conditions of *Ea1* and *Eb1* for $V_{k,m} = 1$ and $W_{k,m} = 1$, respectively).

IV. Asymmetric Game

In this section, we focus on a special case with two different groups of players. As mentioned earlier, we can consider that one group of users are close to the BS and the other group of users are far away from the BS.

A. A General-Sum Game with Two Players

As shown in (11) and (13), the average rewards are (non-linear) functions of $\mathbf{x}_{-k}$. Thus, the average payoff for each action of player k can be given by

$$\begin{aligned}\bar{R}_{k;\mathsf{H},m}(\mathbf{x}_{-k}) &= \mathbb{E}[V_{k,m}] - C_{\mathsf{H}}\\ \bar{R}_{k;\mathsf{L},m}(\mathbf{x}_{-k}) &= \mathbb{E}[W_{k,m}] - C_{\mathsf{L}}\\ \bar{R}_{k;0}(\mathbf{x}_{-k}) &= C_0.\end{aligned}$$

For given $\mathbf{x}_{-k}$, the best response is

$$\mathrm{BR}(\mathbf{x}_{-k}) = \underset{\mathbf{x}_k \in \mathcal{X}}{\operatorname{argmax}}\, u_k(\mathbf{x}_k, \mathbf{x}_{-k}), \tag{14}$$

where $u_k(\mathbf{x}_k, \mathbf{x}_{-k})$ is the (average) payoff when player k chooses the mixed strategy $\mathbf{x}_k$ when the others' mixed strategies are $\mathbf{x}_{-k}$, which is given by

$$u_k(\mathbf{x}_k, \mathbf{x}_{-k}) = \mathbf{x}_k^{\mathrm{T}} \mathbf{d}_k(\mathbf{x}_{-k}). \tag{15}$$

Here,

$$\mathbf{d}_k = [\bar{R}_{k;\mathsf{H},1} \ldots \bar{R}_{k;\mathsf{H},M}\ \bar{R}_{k;\mathsf{L},1} \ldots \bar{R}_{k;\mathsf{L},M}\ \bar{R}_{k;0}]^{\mathrm{T}}, \tag{16}$$

where $\mathbf{x}_{-k}$ is omitted for brevity. In addition, the mixed strategies, denoted by $\mathbf{x}_k^*$, are a NE if

$$u_k(\mathbf{x}_k^*, \mathbf{x}_{-k}^*) \ge u_k(\mathbf{x}_k, \mathbf{x}_{-k}^*), \forall \mathbf{x}_k \in \mathcal{X},\ k = 1, \ldots, K. \tag{17}$$

If the number of users is two, i.e., $K = 2$, the game with the payoff in (15) becomes a general-sum game. In particular, it can be shown that

$$\begin{aligned}\mathbf{u}_1(\mathbf{x}_1, \mathbf{x}_2) &= \mathbf{x}_1^{\mathrm{T}} \mathbf{A}_1 \mathbf{x}_2\\ \mathbf{u}_2(\mathbf{x}_2, \mathbf{x}_1) &= \mathbf{x}_2^{\mathrm{T}} \mathbf{A}_2 \mathbf{x}_1,\end{aligned} \tag{18}$$

where $\mathbf{A}_k$, $k = 1, 2$, are independent of $\mathbf{x}_k$. For simplicity, consider that $M = 1$.

$$\mathbf{A}_1 = \begin{bmatrix} 0 & \frac{e^{-\beta_{1;\mathsf{H}} N_0}}{1+\beta_{1;\mathsf{H}} P_{\mathsf{L}} \alpha_2} & e^{-\beta_{1;\mathsf{H}} N_0} \\ \theta_{1,2} e^{-\beta_{1;\mathsf{L}} N_0} & 0 & e^{-\beta_{1;\mathsf{L}} N_0} \\ 0 & 0 & 0 \end{bmatrix} - \mathbf{c}\mathbf{1}^{\mathrm{T}},$$

where $\mathbf{c} = [C_{\mathsf{H}}\ C_{\mathsf{L}}\ C_0]^{\mathrm{T}}$ and $\mathbf{1}$ is a vector of all 1's. Likewise, we can find $\mathbf{A}_2$. In general, if $\alpha_1 \neq \alpha_2$, the resulting game becomes a two-person general-sum game with bimatrix $(\mathbf{A}_1, \mathbf{A}_2^{\mathrm{T}})$.

B. Extension to Two Groups of Players

In this subsection, we consider the case that players or users can be divided into two groups depending on their distances from the BS located at the center of a cell. When a disc with radius D is used to model a cell, we can consider an inner disc with radius $\bar{D}$ ($< D$). The near-group is the set of users located within the inner disc and the far-group is the set of the other users.

To apply the general-sum game with bimatrix $(\mathbf{A}_1, \mathbf{A}_2^{\mathrm{T}})$ obtained with $K = 2$, we consider the notion of evolutionary game [23] with two different populations. Consider two evolutionary games with the two payoff tables $\mathbf{A}_1$ and $\mathbf{A}_2$ that are not coupled. For symmetric games (i.e., $\mathbf{A}_1 = \mathbf{A}_2$), replicator dynamics can be used to find the evolutionary stable strategy (ESS), which is also a NE [17]. That is, with $\mathbf{A} = \mathbf{A}_1 = \mathbf{A}_2$, the following replicator dynamics can converge to an ESS:

$$\frac{dx_i}{dt} = x_i[(\mathbf{A}\mathbf{x})_i - \mathbf{x}^{\mathrm{T}} \mathbf{A}\mathbf{x}]. \tag{19}$$

In [18], with two different groups (populations) of players with $\mathbf{A}_1 \neq \mathbf{A}_2$, under certain conditions, it is shown that the NE of asymmetric games can be found. In particular, if

$$\mathbf{x}_1^* \in \mathsf{NE}(\mathbf{A}_2) \text{ and } \mathbf{x}_2^* \in \mathsf{NE}(\mathbf{A}_1), \tag{20}$$

where $\mathsf{NE}(\mathbf{A})$ is the set of NE of a two-person symmetric game with a payoff matrix $\mathbf{A}$, and $\mathbf{x}_k^*$ has the same support, $(\mathbf{x}_1^*, \mathbf{x}_2^*)$ becomes an NE of the general-sum game with a bimatrix $(\mathbf{A}_1, \mathbf{A}_2^{\mathrm{T}})$. As a result, using the replicator dynamics[1] for each symmetric game, we can find the NE of the general-sum game with a bimatrix $(\mathbf{A}_1, \mathbf{A}_2^{\mathrm{T}})$. Note that in order to hold this result, it has to be assumed that the supports of $\mathbf{x}_k^*$, $k = 1, 2$, should be the same.

V. Numerical Results

In this section, we present numerical results with $M = 1$ to see how the users in different groups interact. To find the NE, replicator dynamics is used. The user index $k = 1$ is assigned to the near-user and $k = 2$ to the far-user.

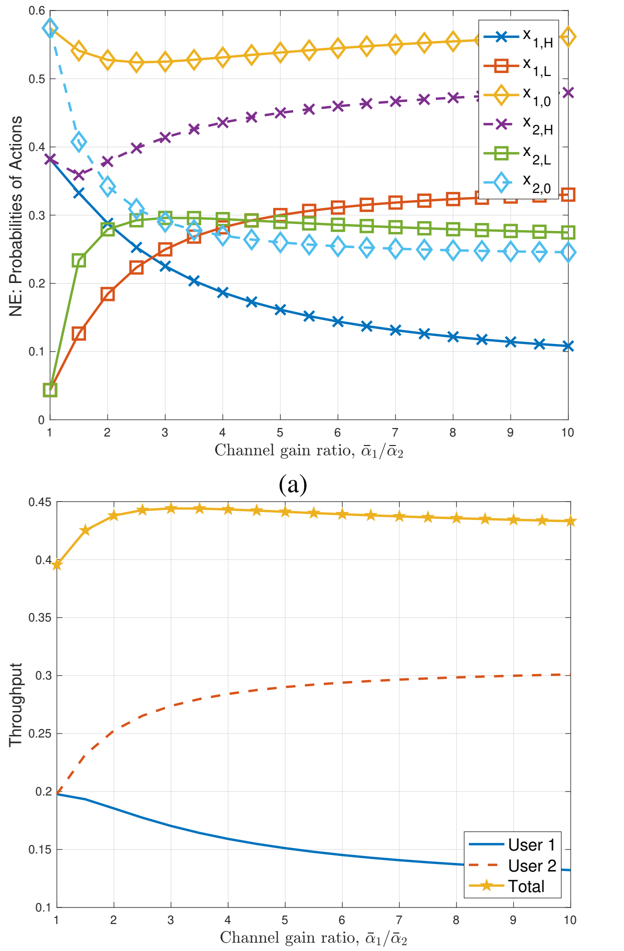
In Fig. 2 (a), with $\Gamma_{\mathsf{H}} = \Gamma_{\mathsf{L}} = 4$ (or 6 dB), $P_{\mathsf{H}} = \Gamma_{\mathsf{L}}(1+\Gamma_{\mathsf{L}})$, $P_{\mathsf{L}} = \Gamma_{\mathsf{L}}$, and $\mathbf{c} = [0.5\ 0.25\ 0.01]^{\mathrm{T}}$, the mixed NE is shown when $\bar{\alpha}_1$ increases, while $\bar{\alpha}_2 = 1$ is fixed. Note that user 1 is the near-user, we expect that $\bar{\alpha}_1$ is larger than $\bar{\alpha}_2$. It is also observed that $x_{1,\mathsf{H}}$ (i.e., the probability that the near-user chooses the high transmit power) decreases as the channel gain, $\bar{\alpha}_1$, increases, while $x_{2,\mathsf{H}}$ increases.

In Fig. 2 (b), we also show the throughput, which is

$$\eta_k = \mathbb{E}_{\sim \mathbf{x}_{-k}^*}[V_k] x_{k;\mathsf{H}}^* + \mathbb{E}_{\sim \mathbf{x}_{-k}^*}[W_k] x_{k;\mathsf{L}}^*, \tag{21}$$

where $\mathbb{E}_{\sim p}$ stands for the expectation over distribution p. That is, the throughput is the average number of successfully transmitted packets with the mixed NE. It is interesting to see that the throughput of user 2 increases as the channel gain of user 1, $\bar{\alpha}_1$, increases. That is, the opponent can exploit a user's improved channel gain when the NOMA-ALOHA system operates at an NE as shown in (20). This behavior is explained in [24], where the users in each group compete themselves, not the users in the other group. Thus, the increase of the channel gain of the near-user does not necessarily provide the throughput gain of themselves, but the far-user. Another important observation is that there is an optimal $\bar{\alpha}_1$ for given $\bar{\alpha}_2$ (i.e., $\frac{\bar{\alpha}_1}{\bar{\alpha}_2} \approx 3$), which maximizes the total throughput as shown in Fig. 2 (b). That is, if a pair of users whose average channel gain ratio is about 3 can be assigned to a channel, it can locally maximize the throughput. Thus, the mixed strategy NE can be used for deriving user-pairing criteria when $M > 1$.

[1]Note that ESS has no meaning for asymmetric games - see [24]. As a result, the mixed strategy NE obtained by the replicator dynamics with the opponent's $\mathbf{A}_k$ is not ESS.



(a)

(b)

Fig. 2. Mixed strategy NE and throughput of the two-person general-sum game with $(\mathbf{A}_1, \mathbf{A}_2^{\mathrm{T}})$ for different values of channel gain ratio, $\frac{\bar{\alpha}_1}{\bar{\alpha}_2}$ when $\Gamma_{\mathsf{H}} = \Gamma_{\mathsf{L}} = 4$, $P_{\mathsf{H}} = \Gamma_{\mathsf{L}}(1+\Gamma_{\mathsf{L}})$, $P_{\mathsf{L}} = \Gamma_{\mathsf{L}}$, and $\mathbf{c} = [0.5\ 0.25\ 0.01]^{\mathrm{T}}$: (a) Mixed strategy NE; (b) Throughput.

VI. Concluding Remarks

We considered the NOMA-ALOHA system that has the two groups of users to randomly access channels without the channel state information for uplink transmissions. In particular, the users in a cell were divided into the two groups such that one near-group of users are located near the BS while the far-group of users are the set of the other users. We developed the game theoretic model for the NOMA-ALOHA and analyzed the average rewards and payoffs of actions independently made by the users. We found the mixed strategy NE of asymmetric games for the NOMA-ALOHA, applying the general-sum game with the two groups of users. Interestingly, the results showed that a far-user can increase the throughput by exploiting the improved channel gain of a near-user. This work revealed that the random access users in each group compete themselves, not the users in the other group, when the NOMA-ALOHA operates with an NE. As further research topics, we will consider optimal user-pairing to maximize the throughput and channel allocation based on fairness.

References

[1] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-lin, and Z. Wang, "Non-orthogonal multiple access for 5g: solutions, challenges, opportunities, and future research trends," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74–81, 2015.

[2] Z. Ding, Y. Liu, J. Choi, M. Elkashlan, C. L. I, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Communications Magazine*, vol. 55, pp. 185–191, February 2017.

[3] J. Choi, "NOMA: Principles and recent results," in *2017 International Symposium on Wireless Communication Systems (ISWCS)*, pp. 349–354, Aug 2017.

[4] J. Choi, "NOMA-based random access with multichannel ALOHA," *IEEE J. Selected Areas in Communications*, vol. 35, pp. 2736–2743, Dec 2017.

[5] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, "Massive machine-type communications in 5G: physical and MAC-layer solutions," *IEEE Communications Magazine*, vol. 54, pp. 59–65, Sep 2016.

[6] J. Ding, M. Nemati, C. Ranaweera, and J. Choi, "IoT connectivity technologies and applications: A survey," *IEEE Access*, vol. 8, pp. 67646–67673, 2020.

[7] E. Altman and Y. Hayel, "A stochastic evolutionary game approach to energy management in a distributed aloha network," *IEEE INFOCOM*, pp. 51–64, 1988.

[8] A. MacKenzie and S. Wicker, "Selfish users in aloha: a game-theoretic approach," in *IEEE 54th Vehicular Technology Conference. VTC Fall 2001. Proceedings (Cat. No.01CH37211)*, vol. 3, pp. 1354–1357 vol.3, 2001.

[9] K. Cohen and A. Leshem, "Distributed game-theoretic optimization and management of multichannel ALOHA networks," *IEEE/ACM Trans. Networking*, vol. 24, pp. 1718–1731, June 2016.

[10] N. Abramson, "THE ALOHA SYSTEM: Another alternative for computer communications," in *Proceedings of the November 17-19, 1970, Fall Joint Computer Conference*, AFIPS '70 (Fall), (New York, NY, USA), pp. 281–285, ACM, 1970.

[11] J. Choi, "Multichannel NOMA-ALOHA game with fading," *IEEE Trans. Communications*, vol. 66, no. 10, pp. 4997–5007, 2018.

[12] J. Choi and J.-B. Seo, "Evolutionary game for hybrid uplink NOMA with truncated channel inversion power control," *IEEE Trans. Communications*, vol. 67, no. 12, pp. 8655–8665, 2019.

[13] T. Börgers and R. Sarin, "Learning through reinforcement and replicator dynamics," *Journal of Economic Theory*, vol. 77, no. 1, pp. 1–14, 1997.

[14] D. Bloembergen, K. Tuyls, D. Hennes, and M. Kaisers, "Evolutionary dynamics of multi-agent learning: A survey," *J. Artif. Int. Res.*, vol. 53, p. 659–697, May 2015.

[15] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2nd ed., 2018.

[16] E. Nisioti and N. Thomos, "Fast Q-learning for improved finite length performance of irregular repetition slotted ALOHA," *IEEE Trans. Cognitive Communications and Networking*, vol. 6, no. 2, pp. 844–857, 2020.

[17] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*. Cambridge, MA: MIT Press, 1998.

[18] K. Tuyls, J. Pérolat, M. Lanctot, G. Ostrovski, R. Savani, J. Z. Leibo, T. Ord, T. Graepel, and S. Legg, "Symmetric decomposition of asymmetric games," *Scientific Reports*, vol. 8, Jan 2018.

[19] F. Kelly and E. Yudovina, *Stochastic Networks*. Cambridge University Press, 2014.

[20] J. Choi, "On throughput bounds of NOMA-ALOHA," *IEEE Wireless Communications Letters*, pp. 1–1, 2022.

[21] W. Yu, C. H. Foh, A. u. Quddus, Y. Liu, and R. Tafazolli, "Throughput analysis and user barring design for uplink NOMA-enabled random access," *IEEE Trans. Wireless Communications*, pp. 1–1, 2021.

[22] J. Choi and Y. Ko, "Reinforcement learning for NOMA-ALOHA under fading (in preparation)." 2022.

[23] J. Weibull, *Evolutionary Game Theory*. Cambridge, MA, USA: MIT Press, 1995.

[24] H. Gintis, *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Behavior*. Princeton, N.J: Princeton University Press, 2nd ed., 2009.