2020 18th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt – RAWNET Workshop)

Invariant Nash Equilibrium for Large Player Games in Multiple Access Channels

Prashant Narayanan and Lakshmi Narasimhan Theagarajan
Department of EE, Indian Institute of Technology Palakkad, Kerala 678557

***Abstract*—We consider a wireless multiple access channel (MAC) with N users. Associated with each user is their time-varying channel state and a finite-length queue which varies with time. In MAC, a receiver decodes the signals of each user by treating the other users' signals as noise. Each user decides their transmit power and the queue-admission control variable dynamically to maximize their expected throughput without any knowledge of the states and actions of other users. This decision problem is formulated as a Markov game for which we show the existence of equilibrium and an algorithm to compute the equilibrium policies. We show that, when the number of users exceeds a given threshold, the expected throughput of all users at all the equilibria points are the same. Furthermore, we show that the equilibrium policies of the users are invariant as long as the number of users remain above the latter threshold, which is referred to as the infinitely invariant Nash equilibrium (IINE). For the considered system, we prove that IINE exists and show that each user can compute these policies using a sequence of linear programs which does not depend upon the parameters of the other users. We also provide the necessary and sufficient conditions for the existence of IINE. Finally, we validate our analysis using numerical simulations.**

***Index Terms*—Multiple access channel, stochastic games, Nash equilibrium, Markov games, best response algorithm, power control, queue constraints, resource allocation.**

I. Introduction

There has been a tremendous growth of wireless communication systems over the last few years. The success of wireless systems is primarily due to the efficient use of their resources. The users are able to obtain their quality of service efficiently in a time varying radio channel by adjusting their own transmission powers. Distributed control of resources for large number of users is an important area of study as it tries to address the high system complexity.

Non-cooperative game theory serves as a natural tool to design and analyze wireless systems with distributed control of resources [1]. In [2], a distributed resource allocation problem using game theory on the multiple access channel (MAC) is considered. They derived the Nash equilibrium for the problem where each user maximizes their own transmission rate in a selfish manner, while knowing the channel gains of all other users. Scutari et al., [3] [4] analyzed the competitive maximization of mutual information in MAC subject to power constraints. They provided sufficient conditions for the existence of unique Nash equilibrium. In a similar setup, [5] showed that for maximizing the effective capacity of each user, there exists a unique Nash equilibrium. Heikkinen [6] analyzed distributed power control problems via potential games.

In [7], the authors consider a MAC model where each user knows only their own channel gain and only the statistics of the channel gains of other users. The problem is formulated as a Bayesian game, for which they show the existence of a unique Nash equilibrium. Altman et al., [8] studied the problem of maximizing throughput of saturated users (i.e., user who always has a packet to transmit) in a Markov channel model and subject to power constraints. They derived the Nash equilibrium for both the centralized scenario, where the base station chooses the transmission power levels for all users, as well as the decentralized scenario, where each user chooses their own power level based on the condition of its radio channel. In [9], the authors showed the convergence of the iterative algorithm proposed in [8]. Altman et al., [10] considered the problem of maximizing the throughput of competitive users in a distributed manner subject to both power and buffer-queue constraints.

The works considered so far, compute equilibrium policies for games with fixed number of users. As the number of users increases, the corresponding equilibrium policies of users change and the complexity of computing these policies also increases. To overcome these problems, population games [11] modeled the number of users present in the system as being significantly large such that each user acts as a selfish agent playing against a continuum of players. In this model, techniques such as evolutionary dynamics to compute the Nash equilibrium policy of a user can be employed. Using the framework of population games, [12] modeled a mobile cellular system, where users adjust their base station associations and dynamically control their transmitter power to adapt to their time varying radio channels. Another technique to overcome the complexity problems was developed in [13], [14]. Here each user interacts with other players only through their average behavior called the mean field. Note that all the users in these models are considered to be interchangeable [11], [13], [14]. An application of mean field modeling in resource allocation was considered in [15], where each user maximizes their own signal to interference and noise ratio (SINR). The authors showed that this problem, as the number of users tends to infinity, can be modeled as a mean field game. The authors in [9] considered a different approach for large number of users. It was shown that when the number of players in their game exceeds a certain fixed threshold, the Nash equilibrium policies of each user gets fixed and can be precomputed in linear time. The authors refer to such a policy as an infinitely invariant Nash equilibrium (IINE) policy. It was

ISBN 978-3-903176-29-4 © 2020 IFIP

also shown that each user requires no information or feedback from other users to compute these policies.

However the model considered in the previous work does not take into account the rate at which data packets arrive form the higher layers. Hence, using policies which are optimal in the saturated model may result in arbitrarily large waiting times for transmission at each user's transmitter. The long delays produced by these policies can significantly reduce the quality-of-service in the wireless network.

In this paper, to mitigate the effect of the above mentioned problems, we consider a model with a finite buffer for each user. Apart from average power constraints, we also consider an average queue constraint for each user. Thus, we consider the problem of dynamic power allocation and queue-control at each user subject to average power as well as queue constraints. Thus, unlike the previously considered models, the current model is more practical as the user's actions affect their transmission waiting time as well. We model this problem as a constrained Markov decision game with independent state information [16]. A major contribution of this paper is proving the existence of the IINE policies and providing the necessary and sufficient conditions for the existence of the IINE policies in the considered system model. We also describe a method to compute IINE of this game, which has relatively low computational complexity.

Notations: s_i denote an element of the set $\mathcal{S}_i$ which is associated with the ith user, $i = 1, \cdots, N$. The set $\mathcal{S}_{-i}$ denotes the set of all $\mathcal{S}_j, j = 1, \cdots, i-1, i+1, \cdots, N$. The expression $1_{\{.\}}$ denotes the indicator function and $(x)^+ = \max(x, 0)$.

II. System model

We consider a wireless multiple access channel (MAC) with N users and a discrete time system model, where n denotes the index of the time slots. Let $h_i[n]$ denote the channel gain of the ith user in the nth time slot. We assume a slow fading model, i.e., the channel gains remain constant over each time slot. Due to the quantization of the channel state information [8], the channel gains belong to a finite, non-negative, ordered set denoted by $\mathcal{H}_i = \{h_i^0, h_i^1, \cdots, h_i^r\}$, where $|\mathcal{H}_i| = r + 1$. We also assume that the h_i's are stationary and ergodic random process. In the nth time slot, the ith user transmits with power $p_i[n]$ such that $p_i[n] \in \mathcal{P}_i$, where $\mathcal{P}_i = \{p_i^0, p_i^1, \cdots, p_i^l\}$ and $|\mathcal{P}_i| = l+1$. The set $\mathcal{P}$ is obtained by the quantization of the transmit power levels [8] and $p_i^0 = 0$.

The packets to be transmitted by the ith user is generated at the higher layers of the ith user's transmitter. At every time slot, $w_i[n]$ packets are sent by the higher layer to the ith user for transmission; $w_i[n]$'s are i.i.d across time and follow the distribution F_i. The incoming packets are stored in the user's buffer until they are transmitted. The buffer size is finite and given by Q_i. The incoming packets are dropped when the buffer is full. In a time slot, each user's buffer accepts incoming packets depending on the value of the admission control variable $c_i[n]$, i.e., packets are accepted into the buffer when $c_i[n] = 1$ and dropped when $c_i[n] = 0$. We assume that, in a given time slot, all arrivals from the upper layers occur after the transmission in that time slot. In each time slot, a user can transmit at most one packet from their buffer and q_i is the number of packets in the buffer, where $q_i \in \mathcal{Q}_i$ and $\mathcal{Q}_i = \{q_i^0, q_i^1, \cdots, Q_i\}$. The queue process $q_i[n]$ evolves as,

$$q_i[n+1] = \min\left((q_i[n] + c_i[n]w_i[n] - 1_{\{p_i[n]>0\}})^+, Q_i\right). \quad (1)$$

We define the Cartesian product $\mathcal{H}_i$ and $\mathcal{Q}_i$ as the set of states $\mathcal{X}_i := \mathcal{H}_i \times \mathcal{Q}_i$ of the user i and the set of actions of user i as $\mathcal{A}_i := \{0, 1\} \times \mathcal{P}_i$. Any element of these sets $\mathcal{X}_i$ and $\mathcal{A}_i$ are represented as $x_i := (h_i, q_i)$ and $a_i := (c_i, p_i)$, respectively, where $c_i \in \{0, 1\}$ is the admission control variable. Each user has an average power and average queue constraints of $\overline{P_i}$ and $\overline{Q_i}$, respectively. We assume that each user knows their instantaneous channel gain and queue state, but is not aware of the channel gains, queue state and transmit power of other users. The transmissions of each user is decoded at the receiver by treating the signals of the other users as noise. Also, we assume $w_i[n]$ is independent of all $h_i[n]$. When N_0 is the receiver noise variance, the reward function associated with user i is given by,

$$t_i(x_i, a_i) \triangleq \log_2\left(1 + \frac{h_i p_i \cdot 1_{\{q_i>0\}}}{N_0 + \sum_{j=1, j\neq i}^{N} h_j p_j \cdot 1_{\{q_j>0\}}}\right). \quad (2)$$

Given this system setup, we formulate a decision problem in the next section.

III. Problem Formulation

We try to address the following problem: identify the optimal rule for each user to non-cooperatively choose an action, i.e., choice of admission control variable and transmit power, such that their average transmission rate (as defined in (2)) is maximized in the MAC system setup described in the previous section. This problem can be mathematically formulated as follows.

The action taken by a user is given by their stationary policy $u_i(a_i|x_i)$ [16], which represents the conditional probability of using the action $a_i \in \mathcal{A}_i$ at a state $x_i \in \mathcal{X}_i$. Let β_i be initial distribution of the states $\mathcal{X}_i$. The occupation measure [17] can be defined on the set $\mathcal{X}_i \times \mathcal{A}_i$ as,

$$z_i(\beta_i, u_i, x_i, a_i) \triangleq \lim_{T\to\infty} \frac{1}{T} \sum_{n=1}^{T} \Pr(x_i[n] = x_i, a_i[n] = a_i|\beta_i, u_i).$$

For convenience, we shall denote the occupation measure as also $z_i(x_i, a_i)$ - dropping the stationary policy and initial distribution from the notation. It can be verified that the above Markov decision process (MDP) is unichain [17]. For a unichain MDP, the occupation measure $z_i(\beta_i, u_i; x_i, a_i)$ is well defined for a stationary policy u_i and is independent of the initial distribution β_i (Theorem. 4.1, [17]). The occupation measure is related to the corresponding stationary policy as,

$$u_i(a_i|x_i) = \frac{z_i(x_i, a_i)}{\sum_{a_i\in\mathcal{A}_i} z_i(x_i, a_i)},\ a_i \in \mathcal{A}_i,\ x_i \in \mathcal{X}_i. \quad (3)$$

In this work, we use occupation measures and stationary policy interchangeably, as one can be obtained from another. Given the policies z_i, the average rate obtained by user i is given by

$$T_i(z_i, z_{-i}) = \sum_{x_i,a_i} R_i^{z_{-i}}(x_i, a_i) z_i(x_i, a_i), \quad (4)$$

where $R_i^{z_{-i}}(x_i, a_i)$ is the instantaneous rate defined as

$$R_i^{z_{-i}}(x_i, a_i) = \sum_{x_{-i}} \sum_{a_{-i}} t_i(x, a) \prod_{l=1}^{N} z_j(x_j, a_j). \quad (5)$$

Similarly, the average power and average queue length under the policy z_i can be defined as,

$$P_i(z_i) = \sum_{x_i,a_i} p_i \cdot z_i(x_i, a_i),\ Q_i(z_i) = \sum_{x_i,a_i} q_i \cdot z_i(x_i, a_i) \quad (6)$$

Any policy z_i that satisfies the user's queue and transmit power constraints is called a feasible policy. We define the set of feasible policies $\mathcal{Z}_i$ as,

$$\mathcal{Z}_i \triangleq \Bigg\{ z_i(x_i, a_i),\, x_i \in \mathcal{X}_i,\, a_i \in \mathcal{A}_i \Bigg| \sum_{(x_i,a_i)} z_i(x_i, a_i) = 1, \quad (7)$$
$$\sum_{(x_i,a_i)} [1_{\{y_i = x_i\}} - \Pr(y_i|a_i, x_i)] z_i(x_i, a_i) = 0,\ \forall\, y_i \in \mathcal{X}_i,$$
$$P_i(z_i) \le \overline{P}_i,\, Q_i(z_i) \le \overline{Q}_i,\ z_i(x_i, a_i) \ge 0,\ \forall\, (x_i, a_i) \in \mathcal{X}_i \times \mathcal{A}_i \Bigg\}.$$

Each user selects a feasible policy to maximize their average rate defined in (4). A feasible policy $z_i^* \in \mathcal{Z}_i$ of user i is called a best response policy when

$$T_i(z_i^*, z_{-i}) - T_i(z_i, z_{-i}) \ge 0, \forall z_i \in \mathcal{Z}_i. \quad (8)$$

We represent the set of all best response policies as $\mathcal{B}_i(z_{-i})$. We formulate this as a non-cooperative Markov game $(\Gamma_{\mathcal{N}})$ and show the existence of Nash equilibria for this game. The ϵ-Nash equilibrium is defined as in [9]. The existence of Nash equilibria for this game has been proved in [10] and the iterative best response algorithm can be computed as described in [9]. The convergence of this the iterative best response algorithm can be shown as in [9]. However, even for moderate number of users, the iterative best response algorithm becomes computationally unfeasible. In the next section, we overcome this problem through infinitely invariant Nash equilibrium (IINE).

IV. Games with large number of users.

The infinitely invariant Nash equilibrium is defined in [9]. The IINE is the equilibrium policy of each user which is invariant, i.e., remains same, even as the number of users in the system changes, as long as the number of users is above the threshold N^* [9]. In the following discussion, we show the existence of an IINE under the assumption of interchangeability of users [13], [14].

First, we define the set of kth sensitive policies [18] as

$$\mathcal{S}_i^k = \big\{ z_i \in \mathcal{S}_i^{k-1} \,\big|\, l_i^k(z_i) = \max_{z_i \in \mathcal{S}_i^{k-1}} l_i^k(z_i) \big\}, \quad (9)$$

where $\mathcal{S}_i^0 = \mathcal{Z}_i$, $\mathcal{S}_i^k \subseteq \mathcal{S}_i^{k-1}$ and l_i^k is defined as

$$l_i^k(z_i) = \sum_{x_i,a_i} (-1)^{k+1} (h_i p_i)^k z_i(x_i, a_i). \quad (10)$$

The set of infinitely sensitive policies is defined as $\mathcal{S}_i = \cap_{k=1}^{\infty} \mathcal{S}_i^k$. That is, $\mathcal{S}_i = \lim_{k\to\infty} \mathcal{S}_i^k$. Now, we show that the set of all IINE policies of user i is indeed the set $\mathcal{S}_i$. We define two users i and j to be interchangeable when $\overline{P}_i = \overline{P}_j$, $\overline{Q}_i = \overline{Q}_j$, $\mathcal{H}_i = \mathcal{H}_j$, $\mathcal{P}_i = \mathcal{P}_j$, $\mathcal{Q}_i = \mathcal{Q}_j$ and $F_i = F_j$.

Theorem 1 (Necessary and sufficient conditions for existence of IINE)**.** *If the set of all users can be be partitioned into finite number of sets $\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_k$ such that all users in a set $\mathcal{N}_l$ are interchangeable and there exists a policy $z_i \in \mathcal{Z}_i$, such that $l_i^1(z_i) > 0, \forall i$, then all the IINE policies of user i belong to $\mathcal{S}_i$. Conversely, every policy of the set $\mathcal{S}_i$ is an IINE policy.*

Proof. Refer Appendix A □

Note that we can ignore the users with $\max_{z_i} l_i^1(z_i) = 0$, from the system as they do not transmit. Thus, we have shown that there exists an IINE policy when the set $\mathcal{S}_i$ is non-empty. However, to compute $\mathcal{S}_i$, we may have to solve the linear program in (9) infinitely. Next, we show that $\mathcal{S}_i^k$ converges after finite number of iterations and, hence, only a *finite* number of iterations of solving linear programs is required to compute an IINE policy.

Theorem 2 (Existence of IINE)**.** *When the conditions stated in Theorem 1 are true, the set of infinitely sensitive policies $\mathcal{S}_i$ is nonempty. Further, $\mathcal{S}_i = \mathcal{S}_i^M$, where M is the distinct number of elements in $\{h_i p_i \big| h_i \in \mathcal{H}_i\,,\, p_i \in \mathcal{P}_i\}$.*

Proof. Refer Appendix C □

Thus, Theorem 2 shows that we need to only solve $M(<\infty)$ linear programs to compute an IINE policy. The objective functions in (10) are only functions of the users states and actions, they do not depend upon any other parameters. Hence, these policies can be precomputed by each user without knowing any information from other users. When the number of users in the system crosses the threshold N^*, the precomputed policies become the Nash equilibrium policies. Thus, for large number of users, the equilibria policies can be computed easily. This avoids the use of algorithms with prohibitive computational complexity such as the iterative best response computation algorithm.

Remark: When $N \ge N^*$ and a new user joins the system, the new user employs their IINE policy, while the existing users continue to employ their previously computed IINE policies. From the preceding discussion, we can see that these policies continue to constitute an equilibrium for the resulting game of $N+1$ players. Thus, once again, the use of IINE policies significantly reduces the computational complexity.

When the set $\mathcal{S}_i$ contains multiple policies, we show that the average rate achieved by all those policies are equal and so, the user can employ any of those policies interchangeably.

Theorem 3 (Interchangeability of IINE policies.). *If z_i and z_i^* are two IINE policies of the user i, then $T_i(z) = T_i(z^*)$.*

Proof. Refer Appendix C □

V. Numerical Results

In this section, we validate our theoretical results using simulations. We denote the index of the largest transmit power and the largest channel gain by l and r, respectively. We consider the set of channel states and power values for each user i to be the same and is equal to $\{0, \frac{1}{r}, \frac{2}{r}, \cdots, 1\}$ and $\{0, 1, \cdots, l\}$, respectively. We consider a Markov fading model with channel state transition probabilities given by
$P(h_i[n] = 0 \big| h_i[n-1] = 0) = \frac{1}{2}$,
$P(h_i[n] = \frac{1}{r} \big| h_i[n-1] = 0) = \frac{1}{2}$,
$P(h_i[n] = \frac{r-1}{r} \big| h_i[n-1] = 1) = \frac{1}{2}$,
$P(h_i[n] = 1 \big| h_i[n-1] = 1) = \frac{1}{2}$,
$P(h_i[n] = \frac{j-1}{r} \big| h_i[n-1] = \frac{j}{r}) = \frac{1}{3}$,
$P(h_i[n] = \frac{j+1}{r} \big| h_i[n-1] = \frac{j}{r}) = \frac{1}{3}$, and
$P(h_i[n] = \frac{j}{r} \big| h_i[n-1] = \frac{j}{r}) = \frac{1}{2}$, $(1 \le j \le r-1)$. The noise variance was fixed to be 1. The arrival distribution of packets for all users is considered to be Poisson with parameter λ. The power and queue constraint for each user is the same, and is denoted by $\bar{P}$ and $\bar{Q}$, respectively. We simulate seven different scenarios with different system configurations. The parameters considered in these scenarios are listed in Table I. M denotes the number of linear programs required to be solved to compute the IINE policy. The minimum number of users at which the IINE policy becomes an equilibrium policy is given by N^*. We can see that there exists a scenario (e.g., scenario 7) where the value of N^* can be even 1, i.e., the IINE policy is an NE policy even for $N \ge 1$.

TABLE I
SIMULATION PARAMETERS

Scenario	r	l	Q	$\bar{P}$	$\bar{Q}$	λ	M	N^*
1	2	2	1	.50000	.500	.49	4	3
2	2	3	1	.95000	.500	.49	6	3
3	2	3	2	1.5500	1.00	.90	6	3
4	3	3	2	1.2800	.650	.60	7	3
5	3	3	3	2.1000	1.60	1.5	7	4
6	2	3	2	1.5500	.900	1.0	6	2
7	2	3	2	1.7000	.900	1.0	6	1

In Figure 1, we plot the l_2 norm distance between the NE policies and the IINE policy of user 1 as the number of users varies. For each value of N, the best response algorithm is used to compute an NE policy $(z_1(N))$ of user 1 for each scenario. In Figure (1), for different scenarios and values of N, we plot $\|z_1(N) - z_1^*\|_2$. The invariant policy is calculated by solving a sequence of linear programs as given in Theorem (2). From Fig. 1, we observe that the NE policy of user 1 quickly converges to the IINE policy when the number of users exceeds N^*.

In Figure 2, for different number of users, we plot the absolute difference between the time average of the rate of user 1 when all the N users use their NE policies and the time average of the rate of user 1 when all the N users use their IINE policies. Once again, we can see that, when $N \ge N^*$, the equilibrium reward of user 1 is the same as the reward obtained when all the users employ their IINE policy. Note that, in all these scenarios, for $N \ge N^*$, the NE policies have become equal to the IINE policies and hence, their rewards also converge.

In figure 3, we compare the IINE policy against a standard

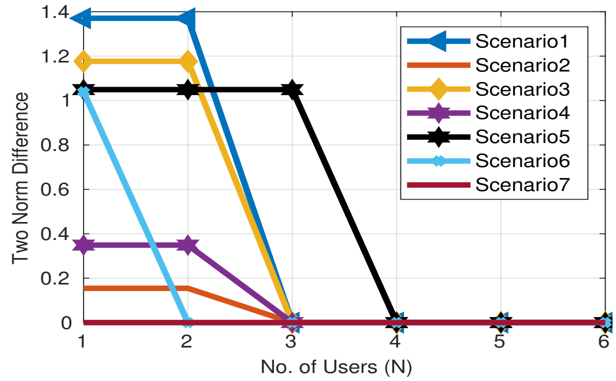

Fig. 1. l_2-norm distance between the NE policy and the IINE policy of the first user for different system configurations and number of users (N).

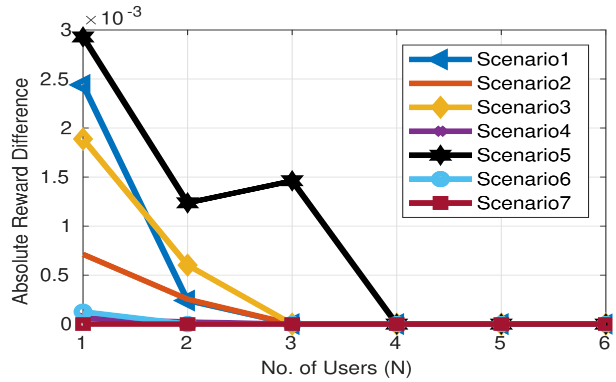

Fig. 2. Absolute difference between the reward of user 1 when all users use their NE policies and the reward of user 1 when all users use their IINE policies for different system configurations and number of users (N).

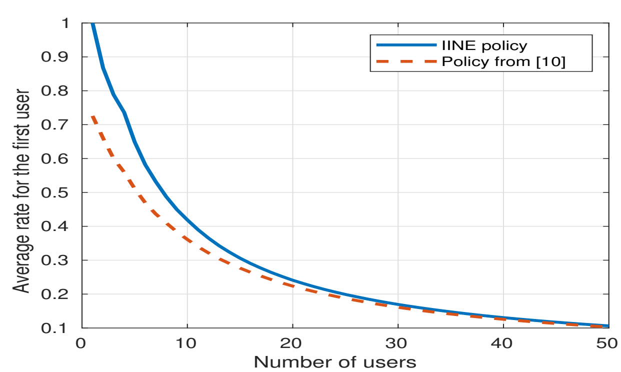

Fig. 3. Comparison of the average rate achieved by the first user for different number of users when the IINE policy and the policy in [10] are employed.

policy proposed in [10], where each user transmits with a probability ρ whenever there is a packet in the buffer. For this simulation, we have set $Q = 5$, $l = 10$, $\lambda = 25$, and $\rho = \frac{1}{2}$. From Fig. 3, we can see that the rate of a user, when all the users employ the IINE policy, outperforms the case when all the users employs the standard policy.

Thus, from the above simulation results we can see that the NE policies of a user converges to the IINE whenever $N \geq N^*$. Further, $\forall N \geq N^*$, one need not use the computationally expensive best response algorithm to compute the NE policy; instead, the IINE policy can be precomputed and employed in all these scenarios.

VI. Conclusions

In this paper, we analyzed the scenario where multiple users, with power constraints and buffer constraints, simultaneously communicate to a single receiver over the multiple access channel. We modeled this problem as a constrained Markov game with independent state information. We proved the existence an infinitely invariant Nash equilibrium (IINE) in this model. We also derived the necessary and sufficient conditions for the existence of the same. We showed that, for the computation of the IINE, the users require only the information of their own states and actions, and does not require any feedback or information from other users in the system. We also showed that an IINE can be computed by solving a finite sequence of Linear programs. Finally, we provided numerical results to validate our analysis.

Appendix A
Proof of theorem 1

The following are some of the notations used in this proof: for two real valued functions $f(n)$ and $g(n)$, the notations $f(n) = o(g(n))$, $f(n) = O(g(n))$ and $f(n) = \Theta(g(n))$ denotes that, there exist constants $c_1 > 0$, $c_2 > 0$ and N_0 such that $\forall n \geq N_0$, $f(n) > c_1 g(n)$, $f(n) < c_2 g(n)$ or $c_1 g(n) < f(n) < c_2 g(n)$, respectively.

Let X_j be a random variable which is defined as

$$\Pr(X_j = h_j p_j) = \sum_{\substack{\forall h_i, p_i \\ \text{s.t. } h_i p_i = h_j p_j}} \sum_{q_i, c_i} z_j(h_i, p_i, c_i, q_i), \qquad (11)$$

$$\text{and} \quad \mu = \inf_{j \geq 0} \mathbb{E}(X_j). \qquad (12)$$

It can be shown that $\mu > 0$ [19]. First, we prove the following lemma which is required to prove Theorem 1.

Lemma 1. *For a natural number k,*

$$\mathbb{E}\left[\frac{1}{\left(\sum_{j\geq 2}^{N+1} X_j + N_0\right)^k}\right] = \Theta\left(\frac{1}{N^k}\right), \mathbb{E}\left[\frac{N^k}{\left(\sum_{j\geq 2}^{N+1} X_j + N_0\right)^k}\right] = \Theta(1). \qquad (13)$$

Proof: We know that $X_j \leq h_j^r p_J^l$, where $h_j^r = \max_i h_j^i$ and $p_j^l = \max_i p_j^i$. Hence, we have

$$\mathbb{E}\left[\frac{1}{\left(\sum_{j\geq 2}^{N+1} X_j + N_0\right)^k}\right] \geq \mathbb{E}\left[\frac{1}{\left(N h_j^r p_j^l + N_0\right)^k}\right] = o\left(\frac{1}{N^k}\right). \qquad (14)$$

Also, for some constant c_1, we have

$$\mathbb{E}\left[\frac{1}{\left(\sum_{j\geq 2}^{N+1} X_j + N_0\right)^k}\right] = \mathbb{E}\left[\frac{1_{\left\{\sum_{j\geq 2}^{N+1} X_j \geq \frac{1}{2} N\mu\right\}} + 1_{\left\{\sum_{j\geq 2}^{N+1} X_j < \frac{1}{2} N\mu\right\}}}{\left(\sum_{j\geq 2}^{N+1} X_j + N_0\right)^k}\right]$$
$$\leq \frac{2^k}{(N\mu)^k} + \frac{1}{N_0^k} P\left(\frac{\sum_{j\geq 2}^{N+1} X_j}{N} < u/2\right)$$
$$\leq \frac{2^k}{(N\mu)^k} + \frac{1}{N_0^k} \exp\left(-c_1 N \mu^2\right) = O\left(\frac{1}{N^k}\right),$$

where the last inequality follows from Hoeffding's inequality. The bounds in (13) follow from the above relationships. ∎

Let z_1^* denote an IINE policy for user i, we shall prove that $z_1^* \in \mathcal{S}_1$ using induction. As z_1^* is an IINE policy, we have from definition [9] and (11) that for any policy $z_1 \in \mathcal{Z}_1$ and for all $N \geq N^*, k > 0$,

$$\sum_{x_1, a_1} \left[\mathbb{E}\left[N^k \log_2\left(1 + \frac{h_1 p_1 1_{\{q_1 > 0\}}}{\sum_{j\geq 2}^{N+1} X_j + N_0}\right)\right] (z_1^*(x_1, a_1) - z_1(x_1, a_1))\right] \geq 0 \qquad (15)$$

For $k = 1$ and some constant $c > 0$, by Taylor series expansion, we get

$$\mathbb{E}\left[N \log_2\left(1 + \frac{h_1 p_1 1_{\{q_1 > 0\}}}{\sum_{j\geq 2}^{N+1} X_j + N_0}\right)\right]$$
$$= \mathbb{E}\left[\frac{N h_1 p_1 1_{\{q_1 > 0\}}}{\sum_{j\geq 2}^{N+1} X_j + N_0}\right] + c\mathbb{E}\left[N \cdot O\left(\frac{1}{\left(\sum_{j\geq 2}^{N+1} X_j + N_0\right)^2}\right)\right]$$
$$\stackrel{(a)}{=} \Theta(1) h_1 p_1 1_{\{q_i > 0\}} + \Theta\left(\frac{1}{N}\right), \qquad (16)$$

where (a) follows from Lemma (1). Let $c_1 > 0$ and $c_2 > 0$; substituting (16) in (15), we get

$$\sum_{x_i, a_i} \left(c_1 h_1 p_1 1_{\{q_1 > 0\}} + \frac{c_2}{N}\right) (z_1^*(x_1, a_1) - z_1(x_1, a_1)) \geq 0.$$

When $N \to \infty$, from the above, we get that $z_i^* \in \mathcal{S}_1^1$.

Now, we shall assume that $z_1^* \in \mathcal{S}_1^m$, $1 < m \leq k-1$ and prove that $z_1^* \in \mathcal{S}_1^k$. From (2) and Taylor series expansion, we have

$$\mathbb{E}\left[N^k \log_2\left(1 + \frac{h_1 p_1 1_{\{q_1 > 0\}}}{\sum_{j\geq 2}^{N+1} X_j + N_0}\right)\right] = c\mathbb{E}\left[-N^k \sum_{l=1}^{k} \left(\frac{-h_1 p_1 1_{\{q_1 > 0\}}}{\sum_{j\geq 2}^{N+1} X_j + N_0}\right)^l\right]$$
$$+ c\mathbb{E}\left[N^k O\left(\frac{1}{\left(\sum_{j\geq 2}^{N+1} X_j + N_0\right)^{k+1}}\right)\right]$$
$$= c\mathbb{E}\left[-N^k \sum_{l=1}^{k-1} \left(\frac{-h_1 p_1 1_{\{q_1 > 0\}}}{\sum_{j\geq 2}^{N+1} X_j + N_0}\right)^l\right] + \Theta(1)\left((-1)^{k+1} h_i p_i 1_{\{q_i > 0\}}\right)^k + \psi, \qquad (17)$$

where the last equality follows from Lemma 1 and $\psi = \Theta\left(\frac{1}{N}\right)$. Substituting (17) in (15), we get

$$\sum_{x_1, a_1} c\mathbb{E}\left[-N^k \sum_{l=1}^{k-1} \left(\frac{-h_1 p_1 1_{\{q_1 > 0\}}}{\sum_{j\geq 2}^{N+1} X_j + N_0}\right)^l\right] (z_1^*(x_1, a_1) - z_1(x_1, a_1)) \qquad (18)$$
$$+ \sum_{x_1, a_1} \left(c_1 \left((-1)^{k+1} h_1 p_1 1_{\{q_1 > 0\}}\right)^k + \frac{c_2}{N}\right) (z_1^*(x_1, a_1) - z_1(x_1, a_1)) \geq 0.$$

For $z_1 \in \mathcal{S}_1^{k-1}$, from (9), we can see that the first term in (18) is 0. Further, as before, when $N \to \infty$, we get that $z_1^* \in \mathcal{S}_1^k$. Hence, by induction, $z_1^* \in \mathcal{S}_1$.

Now, we prove the converse. Without loss of generality we consider user 1 and show that if $z_1^* \in \mathcal{S}_1$, then z_1^* is an IINE policy. We know that $l_1^1(z_1^*) > 0$. Let z_1 denote a feasible policy of user 1 such that it is also a vertex/endpoint of $\mathcal{Z}_i$.

Case (i) $z_1 \notin \mathcal{S}_1$:

Here, we need to prove that there exists a positive number N_{k_1} such that $\forall N \geq N_{k_1}$, $T_1(z_1^*, z_{-1}^*) - T_1(z_1, z_{-1}^*) > 0$. As $z_1 \notin \mathcal{S}_1$, there exists a positive integer k such that $z_1 \notin \mathcal{S}_1^k$ and $z_1 \in \mathcal{S}_1^m$, $1 \leq m \leq k-1$. Let k be the smallest such integer. Now, we have

$$
\begin{aligned}
&(T_1(z_1^*, z_{-1}^*) - T_1(z_1, z_{-1}^*))\, N^k \qquad (19)\\
=& \sum_{x_1, a_1} \left[\mathbb{E}\left[N^k \log_2 \left(1 + \frac{h_1 p_1 1_{\{q_1 > 0\}}}{\sum_{j \geq 2}^{N+1} X_j + N_0} \right) \right] (z_1^*(x_1, a_1) - z_1(x_1, a_1)) \right]
\end{aligned}
$$

For some constant c, using Taylor series expansion of the logarithm term in (19), we get $T_1(z_1^*, z_{-1}^*) - T_1(z_1, z_{-1}^*) = \beta_k + \beta$, where

$$
\begin{aligned}
\beta_k =& \sum_{i=1}^{k} \mathbb{E}\left[\frac{N^k}{\left(\sum_{j \geq 2}^{N+1} X_j + N_0\right)^i} \right] \left[\sum_{x_1, a_1} (-1)^{i+1} \left(h_1 p_1 1_{\{q_1 > 0\}} \right)^i \right.\\
&\left. \cdot (z_1^*(x_1, a_1) - z_1(x_1, a_1)) \right], \qquad (20)\\
\beta =& c\mathbb{E}\left[\frac{N^k}{\left(\sum_{j \geq 2}^{N+1} X_j + N_0\right)^{k+1}} \right]\\
&\cdot \sum_{x_1, a_1} (-1)^k \left(h_1 p_1 1_{\{q_1 > 0\}} \right)^{k+1} (z_1^*(x_1, a_1) - z_1(x_1, a_1)). \qquad (21)
\end{aligned}
$$

Since $z_1 \notin \mathcal{S}_1^k$, $z_1 \in \mathcal{S}_1^m$, $1 \leq m \leq k-1$ and from Lemma 1, we can see that $\beta_k = \Theta(1)$, which we denote by the constant $c_1 (> 0)$. Also, due to Lemma 1, $\beta = \Theta(\frac{1}{N})$, i.e., $\beta = \frac{c_2}{N}$ for some constant $c_2 (> 0)$. Thus, we have $T_1(z_1^*, z_{-1}^*) - T_1(z_1, z_{-1}^*) > c_1 - \frac{c_2}{N}$. Hence there exists a positive number N_1 such that $\forall N \geq N_1$, $T_1(z_1^*, z_{-1}^*) - T_1(z_{1,j}, z_{-1}^*) > 0$. Similarly, there exists N_j for each user j. Let $N_1^* = \max_j N_j$. Therefore, $\forall N \geq N_1^*$, $T_1(z_1^*, z_{-1}^*) - T_1(z_j, z_{-1}^*) > 0$.

Case (ii) $z_1 \in \mathcal{S}_1$:

From (9), $\forall k, l_1^k(z_1^*) = l_1^k(z_1)$. From (11), we obtain two distributions P_* and P_1 corresponding to z_1^* and z_1, respectively. From (10), we can see that same moments of the random variables distributed as P_* and P_1 are the same. Hence, by the method of moments, we have $P_* = P_1$. Thus, $T_1(z_1^*, z_{-1}^*) = T_1(z_{1,}, z_{-1}^*)$.

Therefore, $\forall N \geq N_1^*$ and $\forall z_1 \in \mathcal{Z}_1$, we have $T_1(z_1^*, z_{-1}^*) \geq T_1(z_{1,}, z_{-1}^*)$; hence, z_1^* is an IINE policy. Generalizing this to all users, we get $N^* = \sup\{N_i^*\}$. The quantity N^* exists and is finite due to the assumption of interchangeability of users. ☐

APPENDIX B
PROOF OF THEOREM 2

Given the conditions in Theorem 1, we can see that $\mathcal{Z}_i$ is non-empty. Therefore, by construction (9), we can see that $\mathcal{S}_i^1$ contains at least one element; in general, $\mathcal{S}_i^k$, for all k, contains at least one element as long as $\mathcal{Z}_i$ is non-empty and $\mathcal{S}_i^k$ is a sequence of non-increasing sets, i.e. $\mathcal{S}_i^k \supseteq \mathcal{S}_i^{k+1} \supseteq \mathcal{S}_i$.

Now, we show that $\mathcal{S}_i = \mathcal{S}_i^M$, where M is the number of distinct elements in the set $\{h_i p_i | h_i \in \mathcal{H}_i, p_i \in \mathcal{P}_i\}$. Let z_i and $\hat{z}_i$ be two distinct policies belonging to $\mathcal{S}_i^k$ and $\mathcal{S}_i$, respectively; hence, we have $z_i, \hat{z}_i \in \mathcal{S}_i^k$. We order the values in the set $\{h_i p_i | h_i \in \mathcal{H}_i, p_i \in \mathcal{P}_i\}$ as $\{x_1, x_2, \cdots, x_M\}$, with $x_i \leq x_{i+1}$, $x_1 = 0$ and $x_M = h_i^r p_i^l$. Using (11), we obtain two distributions $\hat{P}$ and P corresponding to $\hat{z}_i$ and z_i, respectively.

Let m_k and $\hat{m}_d$ denote the dth moments of random variables distributed as P and $\hat{P}$, respectively. As $z_i, \hat{z}_i \in \mathcal{S}_i^k$, from (10), we have $m_d = \hat{m}_d$, $1 \leq d \leq M$. Now, we define a matrix $\mathbf{V}$ of size $(M-1 \times M-1)$ with entries $V_{a,b} = (x_a)^b$, $2 \leq a \leq M-1$ and $2 \leq b \leq M-1$. Note that $p_i^0 = 0, \forall i$. Let $\mathbf{z}$ and $\hat{\mathbf{z}}$ represent the vectors containing the probability values of the policies z and $\hat{z}$, respectively, i.e., $\hat{z} = [\hat{P}(x_2), \hat{P}(x_3), \cdots, \hat{P}(x_M)]^T$ and $z = [P(x_2), P(x_3), \cdots, P(x_M)]^T$. Now, we have $\mathbf{Vz} - \mathbf{V}\hat{\mathbf{z}} = 0$. Since $\mathbf{V}$ is an invertible Vandermonde matrix, we have $\hat{\mathbf{z}} = \mathbf{z}$. Hence, from (10), we have $l_i^k(z_i) = l_i^k(\hat{z}_i)$ for all k, i.e., $z_i \in \mathcal{S}_i$ and, hence, $\mathcal{S}_i^k \subseteq \mathcal{S}_i$. Thus, $\forall k \geq M, \mathcal{S}_i^k = \mathcal{S}_i$. ☐

APPENDIX C
PROOF OF THEOREM 3

This proof follows from the proof discussed in Case (ii) of Appendix A.

REFERENCES

[1] Z. Han, D. Niyato, W. Saad, T. Basar, and A. Hjorungnes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*. Cambridge University Press, 2011.

[2] L. Lai and H. E. Gamal, "The water-filling game in fading multiple-access channels," *IEEE Transaction on Information Theory*, vol. 54, no. 5, pp. 2110–2122, May 2008.

[3] G. Scutari, D. P. Palomar, and S. Barbarossa, "Optimal linear precoding strategies for wideband noncooperative systems based on game theory part 1: Nash equilibria," *IEEE Transaction on Signal Processing*, vol. 56, no. 3, pp. 1230–1249, March 2008.

[4] G. Scutari, D. P. Palomar, and S. Barbarossa, "Optimal linear precoding strategies for wideband noncooperative systems based on game theory part 2: Algorithms," *IEEE Transaction on Signal Processing*, vol. 56, no. 3, pp. 1250–1267, March 2008.

[5] D. Qiao, M. C. Gursoy, and S. Velipasalar, "A noncooperative power control game in multiple-access fading channels with qos constraints," *2010 IEEE Wireless Communication and Networking Conference*, pp. 1–6, April 2010.

[6] T. Heikkinen, "A potential game approach to distributed power control and scheduling," *Computer Networks*, vol. 50, no. 13, pp. 2295–2311, September 2006.

[7] G. Heand, M. Debbah, and E. Altman, "A bayesian game-theoretic approach for distributed resource allocation in fading multiple access channels," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, p. 391684, Apr 2010.

[8] E. Altman, K. Avrachenkov, I. Menache, G. Miller, B. J. Prabhu, and A. Shwartz, "Dynamic discrete power control in cellular networks," *IEEE Transactions on Automatic Control*, vol. 54, pp. 2328–2340, Oct 2009.

[9] P. Narayanan and L. N. Theagarajan, "The invariant nash equilibrium for stochastic games in multiple access channel," *2018 16th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, pp. 1–6, May 2018.

[10] E. Altman, K. Avratchenkov, N. Bonneau, M. Debbah, R. El-Azouzi, and D. S. Menasche, "Constrained stochastic games in wireless networks," *IEEE GLOBECOM 2007 - IEEE Global Telecommunications Conference*, pp. 315–320, Nov 2007.

[11] W. H. Sandholm, *Population games and evolutionary dynamics*. MIT press, 2010.

[12] C. Singh, A. Kumar, and R. Sundaresan, “Uplink power control and base station association in multichannel cellular networks,” pp. 43–51, 2009.

[13] O. Gueant, J.-M. Lasry, and P.-L. Lions, “Mean field games and applications,” *Paris-Princeton lectures on mathematical finance 2010*, pp. 205–266, 2011.

[14] J.-M. Lasry and P.-L. Lions, “Mean field games,” *Japanese journal of mathematics*, vol. 2, no. 1, pp. 229–260, 2007.

[15] F. Meriaux, V. Varma, and S. Lasaulce, “Mean field energy games in wireless networks,” *2012 conference record of the forty sixth Asilomar conference on signals, systems and computers (ASILOMAR)*, pp. 671–675, 2012.

[16] E. Altman, K. Avrachenkov, N. Bonneau, M. Debbah, R. El-Azouzi, and D. S. Menasche, “Constrained cost-coupled stochastic games with independent state processes,” *Operations Research Letters*, vol. 36, no. 2, pp. 160 – 164, 2008.

[17] E. Altman, *Constrained Markov decision processes*, vol. 7. CRC Press, 1999.

[18] D. Blackwell, “Discrete dynamic programming,” *The Annals of Mathematical Statistics*, vol. 33, no. 2, pp. 719–726, 1962.

[19] P. Narayanan and L. N. Theagarajan, “Discrete power control stochastic game on the multiple access channel,” *arXiv preprint arXiv:1710.08800*, 2017.