Game Theoretic Cross-Layer Transmission Policies in Multipacket Reception Wireless Networks

Minh Hanh Ngo, Student Member, IEEE, and Vikram Krishnamurthy, Fellow, IEEE

Abstract—We study the structure of the optimal transmission policies for noncooperating nodes in a finite-size random access wireless network, where the medium access control (MAC) protocol is a variant of the time-slotted ALOHA protocol. It is assumed that the network has the multipacket reception capability and every node knows its channel state information (CSI), which is continuously distributed, perfectly at the beginning of each transmission time slot. The objective of each node in the network is to find a transmission policy mapping CSI to transmission probabilities to maximize its individual utility. The problem is formulated as a noncooperative game of a finite number of rational players and actions with a continuous channel state space. We prove that if the probability of success of a node is a nondecreasing function of its CSI, there exists a threshold transmission policy that maximizes its utility. It is then shown that there exists a Nash equilibrium at which every node adopts a threshold policy. The optimality of threshold policies strongly simplifies the problem of optimizing the transmission policy for a node. We propose a stochastic-gradient-based algorithm that exhibits the best response dynamic adjustment process for the transmission game. The theoretical results of the paper as well as the performance of the proposed algorithm are illustrated via numerical examples.

Index Terms—Best response dynamics, channel state information (CSI), multipacket reception, Nash equilibrium, threshold transmission policy.

I. INTRODUCTION

CLASSICAL random access protocols, including ALOHA, are designed based on the idealized collision channel model: if only one node transmits, its packet is received correctly with certainty but if more than one node transmit at the same time, all packets are lost due to collision. However, the collision channel model does not hold in many important practical communication systems. For example, code-division multiple-access (CDMA) systems or systems with multiple antennas at the base station allow one or more packets to be received correctly in the presence of simultaneous transmissions [1].

In [2], the multipacket reception (MPR) model was proposed. The MPR model allows modelling systems where one or more packets can be received correctly with fixed probabilities when multiple nodes transmit simultaneously. A limitation of the MPR model is that channel states do not affect the reception of packets directly and all nodes are indistinguishable. In [3]–[5], it is shown that using the MPR model, a nonzero asymptotic system throughput can be obtained. In addition, a decentralized transmission control algorithm that achieves the best asymptotic system throughput was proposed in [3]. A generalization of MPR to the asymmetrical model is given in [6].

The focus of [7]–[9] is on time-slotted ALOHA systems with selfish nodes that are allowed to select their own transmission policies, which map the numbers of nodes contending the channel to transmission probabilities, to maximize their individual utilities using a game theoretic approach. The existence of a symmetric Nash equilibrium is proved for the collision channel model in [7] and [10] and the MPR reception model in [9].

In [11], the generalized multipacket reception (G-MPR) model was proposed. In the G-MPR model, the probability of receiving a packet correctly depends on the channel states of the transmitting nodes. Hence, the G-MPR model includes the MPR model of [3] as a special case. The G-MPR model provides a framework for exploiting channel state information (CSI) for optimal power or transmission control. In other words, the G-MPR model is a reception model that is suitable for exploiting information from the physical (PHY) layer for medium access control (MAC) layer protocol design.

Early work on exploiting CSI includes [12]–[14], which considered exploiting multiser diversity in the collision channel model via a variant of the ALOHA protocol, namely the channel-aware ALOHA protocol. Using the G-MPR model, [11] proposed a variant of the ALOHA protocol where transmission probability is allowed to be a function of CSI. This is the MAC protocol we consider in this paper. In [11], the problem of optimal transmission control for the spatially homogeneous slotted ALOHA network where all nodes deploy the same transmit probability function is formulated. The structure of the optimal transmission policies for spatially heterogeneous and homogeneous slotted ALOHA networks is studied in [15].

In this paper, the problem of optimal decentralized transmission control is formulated as a noncooperative transmission game and the structure of the optimal transmission policy is studied. The main difference between the work of this paper and early work on the application of game theory to ALOHA networks [7]–[10] is the exploitation of CSI via the G-MPR model and the relaxation of the assumption that all nodes are symmetric. In comparison, the main difference between this paper and other work on exploiting CSI for optimal transmission control [11], [15] is the formulation of the problem as a noncooperative game as well as the introduction of the
transmission and waiting costs. The transmission and waiting costs are necessary for the formulation of the game but they also offer a means to take into account factors such as battery constraints and performance requirements. The waiting cost can be adjusted to enhance long-term fairness in the network and can be increased to reduce transmission delays.

The main results of this paper include the following.

1) We prove that if the probability of correct reception of a packet from a node given the transmission policies of other nodes is a nondecreasing Lebesgue measurable function of its CSI, there exists a threshold transmission policy that maximizes the expected reward of the node. [See (1) for the definition of a transmission policy, and see Fig. 1 for the structure of the optimal policy.]

2) Assuming the nodes select their transmission policies in the set of Lebesgue measurable functions, we proved that the noncooperative transmission game has at least a Nash equilibrium at which every player (node) deploys a threshold transmission policy. (See Fig. 1.)

3) The symmetric network model where the nodes are equidistant from the base station and have the same transmission and waiting costs is also considered. For this special case, we prove that under verifiable, mild conditions, there exists a symmetric Nash equilibrium profile at which all nodes deploy the same threshold transmission policy. In addition, the symmetric Nash equilibrium transmission threshold is a nondecreasing function of the number of active nodes.

4) We study the best response dynamics algorithm and prove its convergence for two-node transmission games. We explicitly characterize the best response functions for the general multinode transmission game, where channel state is exponentially distributed. This characterization allows us to verify deterministically (but numerically) local asymptotic stability of a Nash equilibrium.

5) At each iteration of the best response dynamics algorithm, a node has to solve the stochastic optimization problem [see (23)] for its best response transmission policy. We propose an algorithm that converges to the best response dynamic adjustment process for the transmission game, where each player updates its policy while keeping fixed the strategies of other players. The core of this algorithm is a stochastic gradient algorithm with a constant step size [see (34)], which can be deployed by any node to adaptively estimate its best response transmission policy without knowing the policies, or the channel distribution functions of other nodes.

The paper is organized as follows. Section II describes the wireless network model, Section III is the formulation of the decentralized transmission control problem as a noncooperative transmission game. Section IV presents the theoretical results on the structure of the optimal transmission policies as well as the existence of a Nash equilibrium. In Section V, the existence of a symmetric Nash equilibrium for the symmetric game is proved. An algorithm for estimating the optimal transmission policy and Nash equilibrium is proposed in Section VI. Section VII contains numerical examples.

II. MULTIPACKET RECEPTION NETWORK MODEL

In this section we define the wireless network and the reception model that are considered in the paper. Section II A describes the wireless random access network model with the MAC protocol being a variant of the time slotted ALOHA protocol, where instead of fixed transmission probabilities, each node in the network has a transmission policy mapping its CSI to transmission probabilities [see (1) for the definition of a transmission policy]. Our network model is similar to the model considered in [11]. The G-MPR model, which was proposed in [11], is mathematically described in the Section II-B.

A. Network Model

Consider a time slotted wireless network of $K$ nodes (e.g., sensors in a wireless sensor network), where $K$ is a finite positive integer. Transmission is synchronized at the beginning of each time slot. We consider the uplink communication channel where the nodes communicate with a common base station. Let $i = 1, 2, \ldots, K$ index the nodes in the network and the random variable $\gamma_i$ denote the channel state of node $i$. Assume that $\gamma_i \in [0, M]$, for some finite $M \in \mathbb{R}_+$, which is the set of nonnegative reals. $M$ can be arbitrarily large so that $[0, M]$ includes the entire range of CSI that is of practical interest. Denote the probability distribution function of $\gamma_i$ by $F_i(\cdot)$. Assume that $F_i(\cdot)$ is continuous for all $i = 1, 2, \ldots, K$. Furthermore, it is assumed that at the beginning of each transmission time slot, node $i$ knows its instantaneous CSI, $\gamma_i$, perfectly. Any parameter that influences the reception of packets can be chosen as channel state, for example, channel gain or position of a node with respect to the base station can represent channel state. If channel gain represents channel state, a node can estimate its individual CSI by measure the strength of a beacon signal, which is broadcasted by the base station to all nodes.

In the network, a node that does not have any packet to transmit is referred to as an inactive node. In contrast, a node with at least one packet to transmit is referred to as an active node. At each time slot, inactive nodes perform no action while active nodes must either transmit or wait, i.e., not transmit. The probability with which an active node transmits is determined by its instantaneous CSI and its transmission policy, which is a function that maps CSI to transmit probabilities.

The reception model of the system, which is the G-MPR model, is described in the next subsection. The key difference between the G-MPR model and the conventional
collision model is the fact that in the G-MPR model, the reception of packets depends on the current channel states [e.g., signal-to-noise ratios (SNRs), distances from the base station] of all transmitting nodes. This is usually the case in CDMA wireless networks. In some cases, it is also possible to abstract the reception of an uplink, where the base station uses multiple antennas, into the G-MPR model [11]. In the G-MPR model, the concept of collision is redundant, even though it includes the collision channel model as a special case.

The objective of each node in the network is to find a transmission policy mapping channel states to transmit probabilities to maximize its individual utility. In other words, each node has to solve a function optimization problem, i.e., an infinite dimensional optimization problem. A function optimization problem needs to be defined over a function space. In this paper we consider transmission policies that are Lebesgue measurable and the definition of a transmission policy is given below.

Consider the normed linear function space $L_\infty[0, M]$, which is the space of all Lebesgue measurable functions defined on $[0, M]$ that are bounded almost everywhere (a.e.). The norm of a function in $L_\infty[0, M]$ is its essential supremum [16]. Let $B_{L_\infty[0, M]}[0, 1]$ be the set of all functions in $L_\infty[0, M]$ that have norms in the range $[0, 1]$. Define a transmission policy to be a function mapping channel states of a node to its transmission probabilities, as follows:

$$p_i(\cdot) : [0, M] \rightarrow [0, 1]$$

for $i \in \{1, \ldots, K\}$. A transmission policy is sometimes referred to as a transmit probability function. We only consider transmission policies that are in $B_{L_\infty[0, M]}[0, 1]$. We now define pure, randomized, and threshold transmission policies.

**Definition 1:** A pure transmission policy is a transmit probability function $p(\cdot) : [0, M] \rightarrow [0, 1]$ such that $p(\gamma) \in \{0, 1\}$ for all $\gamma \in [0, M]$ except for possibly a zero measure set (with respect to the probability measure $F(\cdot)$ of the channel state $\gamma$) of values of $\gamma$.

**Definition 2:** A randomized transmission policy is a transmission policy that is not pure. Equivalently, a randomized policy is a transmit probability function $p(\cdot) : [0, M] \rightarrow [0, 1]$ such that $0 < p(\gamma) < 1$ for some nonzero measure set (with respect to the probability measure $F(\cdot)$ of the channel state) of values of $\gamma$.

**Definition 3:** A threshold transmission policy is a transmission policy $p(\cdot) : [0, M] \rightarrow [0, 1]$ such that

$$p(\gamma) = \begin{cases} 0, & \gamma < \theta \\ 1, & \text{otherwise} \end{cases}$$

for some $\theta \in \mathbb{R}_+$, $0 < \theta \leq M$.

**Notation:** We now define the notation that is used in the paper.

A superscript or a subscript $i$ indicates that the node being referred to is node $i$. In comparison, $\mathcal{I}$ is used to refer to the set of nodes indexed by $\{1, 2, \ldots, K\} = \{i\}$. This notation is standard in game theory [17], [18]. $A_k$ is any unordered set of $k$ integers selected from $1, 2, \ldots, K$, $A_k^c \subseteq \{1, 2, \ldots, K\}$. In the paper, $A_k$ is used to specify the set of all transmitting nodes.

$\gamma_{i}^{A_k} = (\gamma_i : i \in A_k^c)$ is a vector representing channel states of the group of nodes indexed by $A_k^c$.

The expected reward (or utility) of node $i$ is denoted by $T_i(p_i(\cdot), \{p_{-i}(\cdot)\})$, where $p_i(\cdot)$ is the policy played by node $i$ and $\{p_{-i}(\cdot)\}$ denotes the set of transmission policies of all other nodes. Throughout the paper, $I(\cdot)$ is the indicator function and $E_F[\cdot]$ represents the expected value of a random variable with respect to some distribution $F(\cdot)$.

**B. Generalized Multipacket Reception Model**

The G-MPR model, proposed in [11], provides explicit incorporation of CSI into the reception of packets. It is also the reception model considered in [15], [19], and [20].

In the G-MPR model, the outcome of a transmission time slot where $k$ nodes indexed by $A_k^c$ transmit belongs to an event space where each elementary event is represented by a binary $k$-tuple $\Theta_{A_k^c} = (\vartheta_i : i \in A_k^c)$, where $\vartheta_i \in \{0, 1\}$ for each $i \in A_k^c$. $\vartheta_i = 1$ indicates that the packet sent by node $i$ is correctly received and $\vartheta_i = 0$ indicates otherwise.

The reception capability of the system is described by a set of $K$ functions, where the $k$th function $\Phi(\gamma_{i}^{A_k^c} ; \Theta_{A_k^c})$ assigns a probability to the outcome $\Theta_{A_k^c}$ when $k$ nodes indexed by $A_k$ with channel state $\gamma_{i}^{A_k}$ transmit, as follows:

$$\Phi(\gamma_{i}^{A_k} ; \Theta_{A_k}) = \mathbb{P}(\Theta_{A_k} | k \text{ nodes transmit, } \gamma_{i}^{A_k})$$

Equation (3) means that the distribution of the possible outcomes $\{\Theta_{A_k}\}$ is determined by the channel states of the transmitting nodes. Consider the CDMA time-slotted system with matched filter receivers and the signal-to-interference-noise ratio (SINR) threshold reception model as an example. Assuming $\text{SNR}$ represents the channel state, the $k$th function $\Phi(\gamma_{i}^{A_k} ; \Theta_{A_k})$ is given by

$$\Phi(\gamma_{i}^{A_k} , \Theta_{A_k}) = \begin{cases} 1, & \text{if } \Theta_{A_k} = \emptyset \\ 0, & \text{otherwise} \end{cases}$$

where $\emptyset = (\vartheta_i : i \in A_k)$ and

$$\vartheta_i = \mathbb{I} \left( \frac{\gamma_i}{1 + \sum_{j \in A_k^c} \gamma_j/N} > \beta \right)$$

where $\gamma_j$ is the SNR of node $j$, $N$ is the spreading gain and $\beta$ is the quality of service requirement (QoS) parameter. The derivation of the SINR threshold reception model for CDMA systems with linear multiuser detectors is given in [21].

It is assumed that the reception model (3) is symmetric. Mathematically, this can be expressed as

$$\Phi(\gamma_{i}^{A_k} ; \Theta_{A_k}) = \Phi(\text{perm}(\gamma_{i}^{A_k}), \text{perm}(\Theta_{A_k}))$$

for any permutation $\text{perm}$ of a $k$-element vector. This symmetric property is satisfied by the SINR threshold reception model (4) as well as most nontrivial system models. It is also an assumption in [11] and [19].

In the network, the nodes do not cooperate and each node is only concerned about its individual utility. Therefore, given the
III. FORMULATION OF THE DECENTRALIZED OPTIMAL TRANSMISSION CONTROL PROBLEM AS A NONCOOPERATIVE GAME

In this section, we formulate the problem of decentralized optimal transmission control for a random access network of a fixed number of active nodes as a noncooperative transmission game, where each node is selfish and rational. In Section III-A, we define the noncooperative game. The utility function and the optimization problem that must be solved by each node are derived in Section III-B.

A. Formulation of the Noncooperative Optimal Transmission Game

Formally, the problem of optimal decentralized transmission control for the network model defined in Section II can be formulated as a noncooperative game with a continuous state space as follows.
- The set of players \( I = \{1, 2, \ldots, K\} \).
- At each time slot, node \( i \) can choose an action \( a_i \in A_i = \{W, T\} \), where \( W \) means to wait and \( T \) means to transmit. A node can also choose to transmit with some probability.
- A strategy is a transmission policy, defined by (1). Pure and randomized transmission policies are defined in Definitions 1 and 2, respectively. Since the space of pure policies is not finite, the existence of a Nash equilibrium is not straightforward.
- Define a profile to be a set of strategies deployed by all nodes in the network: \( \sigma = \{p_1(\cdot), \ldots, p_K(\cdot)\} \).
- A mathematical expression for the utility function (expected reward) of a node given the policies deployed by other nodes is derived in Section III-B.

B. Utility Function and the Decentralized Optimization Problem

In a transmission time slot, if node \( i \) does not transmit, a waiting cost \( c_i^{(W)} \) is recorded; if it transmits, it has to pay a transmission cost \( c_i^{(T)} \). At the end of a transmission time slot, if a packet is received correctly, the node receives a reward of 1 unit. The instantaneous reward of node \( i \) is then determined as follows:

\[
 r_i = \begin{cases} 
 1 - c_i^{(i)}, & \text{if } a_i = T, \hat{a}_i = 1 \\
 -c_i^{(i)}, & \text{if } a_i = T, \hat{a}_i = 0 \\
 -c_i^{(W)}, & \text{if } a_i = W, \text{ i.e., node } i \text{ did not transmit} 
\end{cases} 
\]

(12)

The condition that ensures a successful transmission is more preferable than no transmission, and no transmission is more preferable than an unsuccessful transmission is

\[
 1 > c_i^{(i)} > c_i^{(W)} > 0, \quad \forall i = 1, 2, \ldots, K. 
\]

(13)
By the symmetric property of the reception model (6), the expected reward of node \( i \), denoted by \( T_i(p_i(\cdot), \{\pi_{-i}(\cdot)\}) \), can be easily derived, as follows:

\[
T_i(p_i(\cdot), \{\pi_{-i}(\cdot)\}) = \int_0^M p_i(\gamma_i) \left[ \prod_{j \in A^K} \left( \int_0^M \left[ \prod_{k=1}^K \left( \sum_{j \in A^K} p_k(\gamma_j) \prod_{j' \neq j} \left( 1 - p_{j'}(\gamma_j) \right) \left( 1 - \psi_i(\gamma_i, \gamma_{A^K \setminus i}) \right) \right] \right) \right] dF_i(\gamma_i) - c^{(i)} \psi_i + c^{(i)}
\]

\[
= \int_0^M p_i(\gamma_i) \left[ \prod_{j \neq i} \left( 1 - p_j(\gamma_j) \right) \left( 1 - \psi_i(\gamma_i, \gamma_{A^K \setminus i}) \right) \right] dF_i(\gamma_i) - c^{(i)} \psi_i + c^{(i)}
\]

\[
= \sup_{p_i(\cdot) \in [0,1]} T_i(p_i(\cdot), \{\pi_{-i}(\cdot)\}) - c^{(i)} \psi_i + c^{(i)}
\]

In the paper, (14) and (17) are used interchangeably as the utility function of node \( i \). The problem of maximizing the utility for node \( i \) can be formulated as

\[
\begin{align*}
T_i(p_i(\cdot), \{\pi_{-i}(\cdot)\}) & = \int_0^M p_i(\gamma_i) \left[ \prod_{j \neq i} \left( 1 - p_j(\gamma_j) \right) \left( 1 - \psi_i(\gamma_i, \gamma_{A^K \setminus i}) \right) \right] \prod_{j \neq i} dF_j(\gamma_j) - c^{(i)} \psi_i + c^{(i)} \\
& = \sup_{p_i(\cdot) \in [0,1]} T_i(p_i(\cdot), \{\pi_{-i}(\cdot)\}) - c^{(i)} \psi_i + c^{(i)}
\end{align*}
\]

where \( T_i(p_i(\cdot), \{\pi_{-i}(\cdot)\}) \) is given by (14) and \( B_{L_{\infty},[0,M]}[0,1] \) is defined in Section II-A.

In the remaining of the paper, we focus on Nash equilibrium profiles. A Nash equilibrium profile is a profile at which no player can benefit by unilaterally deviating from its current policy [17], [18].

**Definition 4:** A profile \( \sigma^* = \{p_1^*, \ldots, p_K^*(\cdot)\} \) is a Nash equilibrium if and only if for all players \( i = 1,2,\ldots,K \) we have

\[
T_i(p_i(\cdot), \{p_{-i}^*(\cdot)\}) \geq T_i(p_i(\cdot), \{p_{-i}^*(\cdot)\}) \forall p_i(\cdot) \in B_{L_{\infty},[0,M]}[0,1].
\]

From the above definition of a Nash equilibrium profile, it is clear that at a Nash equilibrium point (18) must hold for all nodes. One observation that can be made at this point is that the allowance for different channel state distributions, \( F_i(\cdot) \), can lead to unfairness in resource allocation unless the transmission and waiting costs are designed to meet some fairness requirement. The designing of transmission and waiting costs is beyond the scope of our paper. In the next section, we focus on proving the existence of a Nash equilibrium at which every player adopts a threshold transmission policy.

IV. OPTIMALITY OF THRESHOLD POLICIES AND EXISTENCE OF A NASH EQUILIBRIUM

Having formulated the problem of optimal decentralized transmission control as a noncooperative transmission game in the previous section, in this section we study the structure of the Nash equilibrium transmission policies. We follow a common technique in game theory to prove the existence of a structured Nash equilibrium profile. This technique consists of three steps:

1) showing that a particular class of policies is optimal (Theorem 1);
2) proving the existence of a Nash equilibrium when the policy space is restricted to this class of policies (Theorem 2);
3) proving the existence of a Nash equilibrium in the original game by showing that a Nash equilibrium in the game with the restricted policy space is also a Nash equilibrium in the original game (Corollary 2).

Readers are referred to [17] and [18] for examples of the early use of this technique in game theory.

\[\text{The authors thank the anonymous reviewers for very detailed comments on this point.}\]
A. Optimality of Threshold Policies

We prove that the utility of a node can always be maximized by a threshold transmission policy.

Theorem 1: Consider a multipacket reception random access network of $K < \infty$ active nodes where the network and reception models are described in Section II. Consider the noncooperative transmission game formulated in Section III, where the problem of optimizing the utility for node $i = 1, 2, \ldots, K$ is given by (18). Assume the reception model (6) of the network satisfies (8) and (11). There exists a transmit probability function that maximizes node $i$'s expected reward (14) and is a threshold policy

$$ p_i^* = \begin{cases} 1, & \text{if } \Psi_i(\gamma_i, \{p_{-i}(\cdot)\}) - c_i^f + c_{aw}^f > 0 \\ 0, & \text{otherwise} \end{cases} $$

for some $\theta \in [0, M]$.

Proof: The proof of this theorem is an application of the bang-bang principle, presented in [22].

The objective of node $i$ is to maximize its utility, which is given by (17):

$$ T_i(p_i(\cdot), \{p_{-i}(\cdot)\}) = \int_0^M p_i(\gamma_i) \times \left( \Psi_i(\gamma_i, \{p_{-i}(\cdot)\}) - c_i^f + c_{aw}^f \right) dF_i(\gamma_i) - c_i^f $$

It can easily be seen that if $\Psi_i(\gamma_i, \{p_{-i}(\cdot)\})$, defined by (15), is Lebesgue measurable, then

$$ p_i^*(\gamma_i) = \begin{cases} 1, & \text{if } \Psi_i(\gamma_i, \{p_{-i}(\cdot)\}) - c_i^f + c_{aw}^f > 0 \\ 0, & \text{otherwise} \end{cases} $$

is a function in $B_{L_{-}[0,M][0,1]}$, and

$$ T(p_i^*(\cdot), \{p_{-i}(\cdot)\}) = \sup_{p_i^*(\cdot) \in B_{L_{-}[0,M][0,1]}} T(p_i^*(\cdot), \{p_{-i}(\cdot)\}). $$

In other words, if $\Psi_i(\gamma_i, \{p_{-i}(\cdot)\})$ is Lebesgue measurable, then the supremum of $T(p_i^*(\cdot), \{p_{-i}(\cdot)\})$ is attained in $B_{L_{-}[0,M][0,1]}$ at $p_i^*(\gamma_i)$, which is defined by (20). We now prove that $\Psi_i(\gamma_i, \{p_{-i}(\cdot)\})$ is Lebesgue measurable and that $p_i^*(\gamma_i)$, defined by (20), belongs to the class of threshold policies.

The success probability of node $i$ is given by (15), as follows:

$$ \Psi_i(\gamma_i, \{p_{-i}(\cdot)\}) = E_{F_{-i}} \left[ \sum_{k=1}^{K} \sum_{A_k \geq i} \prod_{l \in A_k \setminus \{i\}} p_l(\gamma_l) \times \prod_{j \notin A_k \setminus \{i\}} (1 - p_j(\gamma_j)) \Psi_j(\gamma_j, \{p_{-j}(\cdot)\}) \right]. $$

Due to the Lebesgue measurability of the set of functions $\Psi_j(\gamma_j, \{p_{-j}(\cdot)\})$ and all transmission policies, it is clear that $\Psi_i(\gamma_i, \{p_{-i}(\cdot)\})$ is Lebesgue measurable.

In addition, $\Psi_i(\gamma_i, \{p_{-i}(\cdot)\})$ is nondecreasing in $\gamma_i$ by (8), and $\Psi_i(0, \{p_{-i}(\cdot)\}) = 0$ by (11). The latter implies that $\Psi_i(0, \{p_{-i}(\cdot)\}) - c_i^f + c_{aw}^f < 0$.

If $\Psi_i(M, \{p_{-i}(\cdot)\}) - c_i^f + c_{aw}^f \leq 0$, then it must be the case that $\Psi_i(\gamma_i, \{p_{-i}(\cdot)\}) - c_i^f + c_{aw}^f \leq 0$ for all $\gamma_i \in [0, M]$. The optimal solution is then not to transmit, i.e., $p_i^*(\gamma_i) = 0$ for all $\gamma_i \in [0, M]$. This function is a threshold function with the threshold being equal to $M$.

If $\Psi_i(M, \{p_{-i}(\cdot)\}) - c_i^f + c_{aw}^f > 0$, there must exist a threshold $\theta \in [0, M]$ so that $\Psi_i(\gamma_i, \{p_{-i}(\cdot)\}) - c_i^f + c_{aw}^f < 0$ for all $\gamma_i \leq \theta$ and $\Psi_i(\gamma_i, \{p_{-i}(\cdot)\}) - c_i^f + c_{aw}^f > 0$ for all $\gamma_i > \theta$. Hence, the optimal transmission policy defined by (20) is also a threshold policy.

Remarks:

- It can be noted from the above proof that the optimal transmission threshold for player $i$ is the threshold $\theta_i \in [0, M]$ that satisfies

$$ \Psi_i(\gamma_i, \{p_{-i}(\cdot)\}) - c_i^f + c_{aw}^f \begin{cases} \leq 0, & \forall \gamma_i \leq \theta_i \\ > 0, & \forall \gamma_i > \theta_i \end{cases}. $$

- If the condition (8) is strict, i.e., if (9) holds, then $\Psi_i(\gamma_i, \{p_{-i}(\cdot)\})$ is strictly increasing in $\gamma_i$. In this case, there exists a unique $\theta_i$ that satisfies (21). In other words, the transmission policy defined by (20) is the unique optimal policy, and its optimality is strict. Hence, Theorem 1 can be strengthened that the optimal transmission policy is unique and in the class of threshold policies.

Corollary 1: Consider a time-slotted CDMA network of $K < \infty$ active nodes with the SINR threshold reception model (7) for matched filter receivers. There exists a threshold transmission policy solution to the optimization problem (18) for node $i$ in the network, for any $i \in \{1, 2, \ldots, K\}$.

Proof: The SINR threshold reception model (7) satisfies (8) and (11). The proof then follows from Theorem 1.

B. Existence of a Nash Equilibrium

We reformulate the transmission game allowing only transmission policies in the class of threshold policies and prove the existence of a Nash equilibrium for this case. The existence of a Nash equilibrium at which every node adopts a threshold transmission policy for the original game then follows.

Reformulation of the noncooperative transmission game: In light of Theorem 1, it is clear that in order to maximize the expected reward, it is sufficient for a node to search for its transmission policy in the class of threshold policies. Each threshold transmission policy is completely defined by a single parameter: the CSI threshold beyond which the node transmits with certainty. Assuming every node seeks for a transmission policy in the class of threshold policies, the noncooperative transmission game can be reformulated as follows.

- The set of players $I$ is the set of nodes indexed by $i = 1, 2, \ldots, K$.

Each player $i = 1, 2, \ldots, K$ selects a transmission policy in the class of threshold policies defined by (2). Equivalently, player $i$ selects a threshold $\theta_i$, which determines its transmission policy as

$$ p_i(\gamma_i) = \begin{cases} 0, & \gamma_i \leq \theta_i \\ 1, & \text{otherwise} \end{cases}. $$

The space of pure policies is then the set of possible values of $\theta_i$, which is $[0, M]$.
• The utility function of node $i$ is obtained from (17) by replacing $p_j(\cdot)$ by $I(\gamma > \theta_j)$ for $j = 1, 2, \ldots, K$:

$$T_i(\theta_i, \theta_{-i}) = \int_0^M I(\gamma_i > \theta_i) \times \prod_{j \neq i, j=1}^K \left(1 - I(\gamma_j > \theta_j)\right) \psi_i (\gamma_i, \gamma_{A_k}^i)$$

$$\times \prod_{j \neq i, j=1}^K dF_j(\gamma_j) - e^{(i)} + e^{(j)} dF_i(\gamma_i) - e^{(i)}.$$

Then, the optimization problem (18) can be rewritten as

$$\sup_{\theta_i \in [0,M]} T_i(\theta_i, \theta_{-i})$$

where $T_i(\theta_i, \theta_{-i})$ is given by (22).

Having reformulated the transmission game, in Theorem 2, we use a classical result on the existence of a Nash equilibrium for games with continuous utility functions ([23]–[25], also see [18, Theorem 1.2]) to prove the existence of a Nash equilibrium for the new game.

**Theorem 2:** Consider a multipacket reception random access network of $K < \infty$ active nodes, where the system and reception models are given in Section II. Assume that the nodes only select their transmission policies in the class of threshold transmission policies. Assume that the reception model (6) of the network satisfies (8) and (11). The noncooperative transmission game formulated above, where the problem of optimizing the expected reward for node $i = 1, 2, \ldots, K$ is given by (23), has a Nash equilibrium profile.

**Proof:** See the Appendix.

**Remark:** The assumption that the channel distribution function is continuous is essential in the proof of Theorem 2. If the channel distribution is not continuous it is not guaranteed that the space of pure policy is a compact, convex subset of any Euclidean space and the classical result on the existence of a Nash equilibrium for infinite games with continuous payoff [23]–[25] cannot be used.

Finally, we claim that the strategic noncooperative transmission game formulated in Section III, where the players search for their policies in the whole subset $B_{L_{\infty}[0,M]}[0,1]$ of the normed linear function space $L_{\infty}$, also has a Nash equilibrium profile.

**Corollary 2:** Consider a multipacket reception random access network of $K < \infty$ active nodes, where the system and reception models are given in Section II. Assume that the reception model (6) of the network satisfies (8) and (11). The problem of optimal decentralized transmission control for this network is formulated as a noncooperative transmission game in Section III. This noncooperative transmission game has a Nash equilibrium at which every node deploys a threshold transmission policy.

**Proof:** According to Theorem 2, the transmission game where only threshold transmission policies are considered has a Nash equilibrium. Furthermore, due to the optimality of threshold policies (Theorem 1), a Nash equilibrium profile in the transmission game considering only threshold policies must also be a Nash equilibrium profile in the original transmission game, where a node can use any transmission policy in $B_{L_{\infty}[0,M]}[0,1]$. The existence of a Nash equilibrium at which every node deploys a threshold policy then follows.

**Remark:** If the inequality (8) is strict, i.e., if (9) holds, then threshold transmission policies are strictly optimal (see the remark after the proof of Theorem 1). As a result, all Nash equilibria can be obtained in the class of threshold policies. In comparison, if (9) does not hold, the game may admit a Nash equilibrium outside the class of threshold policies. Furthermore, in this case, it is not immediately clear whether there exists a Nash equilibrium that is Pareto dominant outside the class of threshold policies.

In light of the theoretical results presented in this section, in the remaining of the paper we only consider transmission policies in the class of threshold policies. For convenience, we introduce the definition of a Nash equilibrium transmission threshold vector.

**Definition 5:** A transmission threshold vector $\theta^* = (\theta_1^*, \ldots, \theta_K^*)$ represents a Nash equilibrium if and only if, for all players $i = 1, 2, \ldots, K$,

$$T_i(\theta_i^*, \theta_{-i}^*) \geq T_i(\theta_i, \theta_{-i}) \forall \theta_i \in [0, M].$$

In the next section, we study the existence of a symmetric Nash equilibrium for the symmetric network model where all nodes have the same channel distribution function, transmission and waiting costs.

**V. SYMMETRIC NASH EQUILIBRIUM FOR THE SYMMETRIC GAME**

In this section, we consider the symmetric multipacket reception network where all nodes have the same continuous channel distribution function, i.e., $F_j(\cdot) = F(\cdot) \forall j \in \{1, 2, \ldots, K\}$, the same transmission and waiting costs. In light of the optimality of threshold policies (Theorem 1), let the nodes search for their transmission policies in the class of threshold policies. The existence of a symmetric Nash equilibrium for the symmetric transmission game is proved under a mild condition on the continuity of the success probability function (15) (Theorem 3). It is also shown that the symmetric Nash equilibrium threshold is a nondecreasing function of the number of active nodes in the network (Theorem 4).

We now define the two conditions required by Theorems 3 and 4. Let all nodes except node $i$ deploy the same threshold transmission policy, i.e., $p_j(\gamma_j) = I(\gamma_j > \theta) \forall j \neq i$, for some
\[ \Psi_i(\gamma_i, \theta) = E_{\{\mathcal{F}_{-i}\}} \left[ \sum_{k=1}^{K} \sum_{\mathcal{A}_k^i} \prod_{\ell \neq k} \mathbb{I}(\gamma_\ell > \theta) \times \prod_{j \in \mathcal{A}_k^i} \left(1 - \mathbb{I}(\gamma_j > \theta)\right) \Psi_i \left(\gamma_i, \mathcal{A}_k^i \setminus \{i\}\right) \right] \]  

\[ \text{for } i \in \{1, \ldots, K\}. \]

The first condition of Theorem 3 is that (25) is a continuous function of the CSI \( \gamma_i \). That is

\[ \lim_{\gamma_i \to 0} \Psi_i(\gamma_i, \theta) = \Psi_i(\gamma_0, \theta) \forall i, \gamma_0, \theta \in [0, M], \]  

for \( i \in \{1, \ldots, K\} \).

Writing \( \Psi_i(\gamma_i, \theta) \) in the integral form reveals that the continuity of \( \Psi_i(\gamma_i, \mathcal{A}_k^i \setminus \{i\}) \) with respect to \( \gamma_i \) is a sufficient condition for (26).

The second condition of Theorems 3 and 4 is that

\[ \Psi_i(M, 0) > c_t - c_w \]  

which implies that \( \Psi_i(M, \theta) > c_t - c_w \) for all \( \theta \in [0, M] \) since \( \Psi_i(\gamma_i, \theta) \) is nondecreasing in \( \theta \) and \( \gamma_i \) by conditions (10) and (8), respectively.

Intuitively, (27) means that if the channel state of a node is \( M \), which is the maximum value in the channel state space, and the node transmits, then it will gain a positive expected reward. This condition is not very restrictive due to the fact that \( M \) can be arbitrarily large. In addition, it might be desirable that the transmission and waiting costs are adjusted so that (27) holds in order to eliminate the possibility that the optimal policy for node \( i \) is to never transmit. It should be noted that (27) is sufficient (but not necessary) to ensure that \( \pi(\cdot) = 0 \) is not optimal.

**Theorem 3:** Consider a symmetric multipacket reception random access network of \( K \) active nodes. Assume that (26) and (27) are satisfied. There exists a Nash equilibrium profile at which every node deploys the same threshold transmission policy, i.e., there exists a Nash equilibrium profile \( \sigma^* = \{p_1^*(\cdot), \ldots, p_K^*(\cdot) : p_k^*(\gamma_i) = \mathbb{I}(\gamma_i > \theta_k^*) \forall i = 1, 2, \ldots, K \} \) for some \( \theta^* \in [0, M] \).

**Proof:** See the Appendix.

**Theorem 4:** Consider a symmetric multipacket reception random access network of \( K \) active nodes. Assume that (26) and (27) are satisfied. Let \( \eta \) be a transmission threshold such that \( \sigma^K = \{p_i(\gamma_i) = \mathbb{I}(\gamma_i > \eta) : i = 1, 2, \ldots, K \} \) is a Nash equilibrium profile.

Consider the same network but with \( K+1 \) nodes. There exists \( \xi \geq \eta \) such that \( \sigma^{K+1} = \{p_i(\gamma_i) = \mathbb{I}(\gamma_i > \xi) : i = 1, \ldots, K+1 \} \) is a Nash equilibrium profile for the transmission game of \( K+1 \) players.

**Proof:** See the Appendix.

VI. **STOCHASTIC GRADIENT-BASED ALGORITHM FOR ESTIMATING THE OPTIMAL TRANSMISSION POLICY AND Nash Equilibrium**

The main contribution of our paper has been to prove the optimality of threshold transmission policies and the existence of a Nash equilibrium at which every node adopts a threshold transmission policy. In this section, we are interested in the secondary problem of estimating a Nash equilibrium for the transmission game. Although there exist many learning algorithms in game theory for estimating a Nash equilibrium [26], proving the convergence of these algorithms is very difficult. Here we focus on the best response dynamics algorithm for estimating a Nash equilibrium and exploit our structural result on the optimality of threshold transmission policies. It is known that if the best response dynamics algorithm (Algorithm 1 below) converges, then it converges to a Nash equilibrium [26]. However, apart from the two-user case, we have been unable to give useful sufficient conditions for convergence of Algorithm 1. To give more insight into the sufficient condition for convergence of Algorithm 1 below, we also give an explicit characterization of the best response functions for networks with exponentially distributed channel state (i.e., Rayleigh fading channel).

At each iteration of Algorithm 1 below, it is assumed that each node can solve the stochastic optimization problem (23) for its best response transmission policy. This optimization problem can be solved efficiently using the stochastic gradient ascent method (34) in Section VI-B. Using (34), we propose Algorithm 2 (in Section VI-B) that converges to the best response dynamic adjustment process for the transmission game. In this algorithm, each node uses a gradient estimator and a constant step size to estimate its best response transmission policy via (34). Algorithm 2 can track the slowly time-varying best-response dynamic adjustment process for a slowly time-varying network.

A. **Deterministic Best-Response Dynamics Algorithm for the Transmission Game**

First, we restate the utility optimization problem that needs to be solved numerically by each node. Specifically, in order to maximize its individual expected reward, node \( i \) must solve the optimization problem (23): \( \sup_{\theta_i \in [0, M]} T_i(\theta_i, \theta_-) \).

In light of Theorem 1, the supremum of \( T_i(\theta_i, \theta_-) \) is attainable and it follows from (21) that

\[ H_i(\theta_-) \triangleq \arg \max_{\theta_i \in [0, M]} T_i(\theta_i, \theta_-) = \theta_i \in [0, M] : \]

\[ \Psi_i(\gamma, \theta_-) - c_t^{(i)} + c_w^{(i)} \begin{cases} 
\leq 0, & \forall \gamma \leq \theta_i \\
> 0, & \forall \gamma > \theta_i 
\end{cases} \]  

where \( \Psi_i(\gamma, \theta_-) \) is given by (15) with \( p_j(\cdot) = \mathbb{I}(\gamma \geq \theta_j) \) for all \( j = 1, 2, \ldots, K \). \( H_i(\theta_-) \) defined above is the best-response function for node \( i \).

In this section, we assume (9), which implies that the best response of each player is unique (see the remark after the proof of Theorem 1). \( H_i(\theta_-) \) is then a well-defined scalar value function mapping from \([0, M]^{K-1}\) to \([0, M]\). In addition, \( H_i(\theta_-) \) is nonincreasing in every component of \( \theta_- \) due to (10).
The best-response dynamics algorithm (see [26] for details) for the transmission game, where the nodes take turn to update their transmission policies, can be written as below.

Algorithm 1: Best Response Dynamics of the Transmission Game

- Initialization: At batch \( n = 1 \), select \( \theta_i^{(n)} \in [0, M] \).
- For \( n = 1, 2, \ldots \)
  - For node \( i = 1, 2, \ldots, K \)
    - \( \theta_i^{(n)} = H_i \left( \theta_i^{(n-1)} \right) \), where
      \[
      \theta_i^{(n)} = \begin{cases} 
      \theta_j^{(n)} & j < i, \\
      \theta_i^{(n-1)} & j = i, \\
      \theta_j^{(n)} & j > i.
      \end{cases}
      \] (29)

1) Convergence of Algorithm 1 to a Nash Equilibrium: At this stage, we have not yet given a method to solve (28), we will do so in Section VI-B via a stochastic approximation algorithm that can estimate the optimal \( \theta_i \) for node \( i \). However, temporarily assuming that (28) can be solved, i.e., the mapping (29) can be done in each iteration, it is well known that if Algorithm 1 converges then it always converges to a Nash equilibrium [18]. So we first focus on giving conditions for Algorithm 1 to converge.

First, we present a lemma on global convergence, i.e., convergence for any initial condition, of Algorithm 1 for the two-node case.

**Lemma 1:** For a two-node transmission game, Algorithm 1 converges to a Nash equilibrium.

**Proof:** See the Appendix. ■

Apart from the two-node case, proving global convergence of Algorithm 1 is very difficult; hence, we focus on local convergence conditions. Local asymptotic stability of a Nash equilibrium is sufficient for Algorithm 1 to converge locally, i.e., for initial conditions in some ball around the Nash equilibrium. In Lemma 2, we present a sufficient condition for local asymptotic stability of a Nash equilibrium.

**Lemma 2:** Local Asymptotic Stability of a Nash equilibrium. Let \( \theta^{\ast} \) be a fixed point of Algorithm 1. Assume that \( \Psi_i(\theta_i, \theta_{-i}) \) is differentiable (and hence continuous) at \( \theta^* \) for all \( i = 1, 2, \ldots, K \). Then

\[
\Psi_i(\theta_i^* \theta_{-i}^*) - c_{i,i}^{(i)}(i) = 0, \quad \forall i = 1, 2, \ldots, K
\] (30)

Denote the gradient matrix by \( D(\theta^*) = (d_{i,j})_{i,j=1}^{K} \), where \( d_{i,j} = \nabla_{\theta_i} \Psi_i(\theta_i, \theta_{-i}) \). If the eigenvalues of \( D(\theta^*) \) lie strictly within the unit circle then \( \theta^* \) is a asymptotic stable fixed point of Algorithm 1.

**Proof:** (30) follows from (28) and the assumption on continuity of \( \Psi_i(\cdot, \cdot) \) at the fixed point. It then follows straightforwardly from [27, Theorem C.7.1, p. 334] that if the eigenvalues of \( D(\theta^*) \) lie within the unit circle, then \( \theta^* \) is a asymptotic stable fixed point of Algorithm 1 ■

It is difficult to analytically verify the condition in Lemma 2 since it is not known a priori where the Nash equilibrium is. Later on, we give an explicit representation for the success probability function \( \Psi_i(\cdot, \cdot) \) in (28) and (30), which allows us to numerically evaluate the Nash equilibria and verify local asymptotic stability of a Nash equilibrium.

### B. Explicit Characterization of Best-Response Function for Rayleigh Fading Channels

In this section, we give an explicit representation of the best response functions for CDMA networks with Rayleigh fading channels. This allows us to numerically and deterministically (i.e., without requiring any stochastic simulation) characterize the Nash equilibrium and numerically verify local convergence of Algorithm 1 by studying the deterministic system of (28) for all \( i = 1, 2, \ldots, K \). We assume a Rayleigh channel (i.e., exponentially distributed power) and the SINR threshold reception model (7), where SNR represents channel state and hence can explicitly compute \( \Psi_i(\gamma_i, \theta_{-i}) \) in (28).

Denote the Rayleigh channel probability density function of node \( i \) by \( f_i(\gamma_i) = \lambda_i e^{-\lambda_i \gamma_i} \), where the exponential parameter \( \lambda_i \geq 0 \).

**Lemma 3:** For a symmetric network i.e., all Rayleigh channels are identically distributed \((\lambda_i = \lambda \forall i = 1, 2, \ldots, K)\), the best-response function (28) for node \( i \) has the following representation for \( \Psi_i(\gamma_i, \theta_{-i}) \):

\[
\Psi_i(\gamma_i, \theta_{-i}) = \sum_{k=1}^{K} \sum_{\mathcal{A}_k^i \ni k} \prod_{j \in \mathcal{A}_k^i - i} (1 - e^{-\lambda_i \gamma_i}) \left( a \geq \sum_{j \in \mathcal{A}_k^i - i} \theta_j \right) \times \left( I(k = 1) + I(k \geq 2) \left( e^{-\lambda \sum_{j \in \mathcal{A}_k^i - i} \theta_j} - e^{-\lambda a} \right) + \sum_{l=1}^{k-2} \lambda_l e^{-\lambda_l a} \left( a - \sum_{j \in \mathcal{A}_k^i - i} \theta_j \right) \right) \right) \right) \right) \right)
\] (31)

Here, \( a = N \gamma / \beta - N, \) and \( \beta \) are the spreading gain and the QoS requirement parameter of the SINR threshold reception model (7), respectively.

For an asymmetric network, when the nodes have Rayleigh fading channels with different probability density functions: \( \lambda_i \neq \lambda_j \forall i \neq j, \) \( \Psi_i(\gamma_i, \theta_{-i}) \) in (28) is given by

\[
\Psi_i(\gamma_i, \theta_{-i}) = \sum_{k=1}^{K} \sum_{\mathcal{A}_k^i \ni k} \prod_{j \in \mathcal{A}_k^i - i} (1 - e^{-\lambda_i \gamma_i}) \left( a \geq \sum_{j \in \mathcal{A}_k^i - i} \theta_j \right) \times \left( I(k = 1) + I(k \geq 2) \left( e^{-\lambda \sum_{j \in \mathcal{A}_k^i - i} \theta_j} - \sum_{j \in \mathcal{A}_k^i - i} e^{-\lambda_j a} \right) + I(k \geq 3) e^{-\lambda \sum_{j \in \mathcal{A}_k^i - i} \lambda_j \theta_j} \left( 1 + (-1)^{k-1} \sum_{j \in \mathcal{A}_k^i - i} \left( \prod_{m \in \mathcal{A}_k^i - i - j} \frac{\lambda_m}{\lambda_j - \lambda_m} \right) e^{-\lambda_j \left( a - \sum_{m \in \mathcal{A}_k^i - i - j} \theta_m \right)} \right) \right) \right) \right)
\] (32)
Here, \( a = N\gamma/\beta - N \), \( N \) and \( \beta \) are the spreading gain and the QoS requirement parameter of the SINR threshold reception model (7) respectively.

Proof: See the Appendix.

We will use Lemma 3 to analytically compute the best response functions for a simple three-player transmission game and verify local asymptotic stability of a Nash equilibrium in Section VII-A.

C. Stochastic Gradient Algorithm for Estimating the Optimal Transmission Policy and Nash Equilibrium

In Algorithm 1, it is assumed that at each iteration, node \( i \) can do the mapping (29), i.e., it can solve (23) for its best-response transmission policy. In this section, we propose an algorithm (Algorithm 2 below), where (29) in Algorithm 1 is replaced by a stochastic gradient ascent algorithm that can estimate the best response transmission threshold for any node \( i \), i.e., numerically solve (23). As shown below, the asymptotic behavior of Algorithm 2 is the same as that of Algorithm 1. It should be noted that each node needs to estimate its optimal transmission policy without knowing the policies, the channel distributions of other nodes, or the number of nodes in the network. Hence, the optimization problem (23) is a stochastic optimization problem.

Because the gradient descent is an unconstrained algorithm, an important condition for the validity of Algorithm 2 is that the optimization problem (23) without the constraint \( \theta_l \in [0, M] \) actually has a solution in \([0, M]\), or equivalently, \( \Psi_l(\theta_l, \theta_{-i}) > c_t - c_w \) changes sign from negative to strictly positive at some \( \theta_l^{*} \) in \([0, M]\) for all \( \theta_{-i} \in [0, M] \). A sufficient condition for this fact is

\[
\Psi_l(M, 0) > c_t - c_w \tag{33}
\]

where \( \Psi_l(\theta, \cdot) \) is given by (15) with \( p_j(\cdot) = I(\gamma_j > \theta^*) \), \( j = 1, 2, \ldots, K \). Similar to the explanation of condition (27) for the case of a symmetric network, \( \Psi_l(0, \theta_{-i}) = c_t + c_w = -c_t + c_w < 0 \) and (33) implies that \( \Psi_l(M, \theta_{-i}) = c_t + c_w = 0 \) for all \( \theta_{-i} \in [0, M] \). Hence, (33) is sufficient to guarantee that the solution of the unconstrained version of (23) is in \([0, M]\).

Algorithm 2: Stochastic Best Response Dynamic of the Transmission Game

- Initialization: Outer loop index \( n = 1 \), a batch size \( \Delta \geq 1 \), \( \theta_i \), a constant step size \( 0 < \epsilon \ll 1 \).
- For \( n = 1, 2, \ldots \)
  - For node \( i = 1, 2, \ldots, K \)
    - Node \( i \) updates its transmission threshold once every \( m \) transmission time slots, other nodes keep their policies fixed, i.e., \( \theta_{-i} = ((\theta_{-i}^{(m)})_{j \neq i}, (\theta_{-i}^{(m-1)})_{j \neq i}) \).
  - For Inner loop index \( l = 1, 2, \ldots, \Delta \)
    - Estimate the gradient of \( \Delta \theta_{-i}^{(m)}T_i^{(l)}(\theta_i^{(m)}, \theta_{-i}) \) via a gradient estimator [see (36) and (37)]
    \[
    \theta_i^{(l+1,m)} = \theta_i^{(l,m)} + \epsilon_i \hat{\nabla}_{\theta_i^{(m)}} T_i^{(l)}(\theta_i^{(m)}, \theta_{-i}). \tag{34}
    \]

1) Implementation of Gradient Estimator in (34): First, the utility function \( T_i(\theta_i, \theta_{-i}) \) of node \( i \), given by (22), can be written as

\[
T_i(\theta_i, \theta_{-i}) = \int_0^M I(\gamma_i > \theta_i)(\Psi_i(\gamma_i, \theta_{-i} - c_t^i) + c_t^i) dF_i(\gamma_i) - c_w^i \tag{35}
\]

where \( \Psi_i(\gamma_i, \theta_{-i}) \) is given by (15). \( T_i(\theta_i, \theta_{-i}) \), given by (35), is differentiable almost everywhere by the Fundamental Theorem of Calculus [28]. Furthermore, the gradient of \( T_i(\theta_i, \theta_{-i}) \) is

\[
\nabla_{\theta_i} T_i(\theta_i, \theta_{-i}) = \left( - \frac{\Psi_i(\gamma_i, \theta_{-i}) - c_t^i + c_w^i}{F_i(\gamma_i)} f_i(\gamma_i) \right) f_i(\gamma_i). \tag{36}
\]

Since \( f_i(\gamma_i) \) can be absorbed into the step size, to estimate the gradient of \( T_i(\theta_i, \theta_{-i}) \) we only need to obtain an unbiased estimate of \( \Psi_i(\theta_i, \theta_{-i}) \), which is the probability that the packet from node \( i \) is received correctly given its channel state is equal to \( \theta_i \). If the network is static during the period of \( m \) transmission time slots of the \( l \)th iteration, the unbiased estimate of \( \hat{\Psi}_i^{(l)}(\theta_i^{(l)}, \theta_{-i}) \) is

\[
\hat{\Psi}_i^{(l)}(\theta_i^{(l)}, \theta_{-i}) = \frac{\text{Number of ACK}(\gamma_i = \theta_i)}{\text{Number of Transmissions}(\gamma_i = \theta_i)}. \tag{37}
\]

As an example, consider a system where SNR represents the channel state of a node. The probability of success of node \( i \) when its channel state is \( \theta_i \) can be sampled using a power control mechanism to ensure that the received SNR is equal to \( \theta_i \) and counting the number of ACK(s) that are sent to the node by the station. Hence, a node can obtain the unbiased estimate \( \hat{\Psi}_i^{(l)}(\theta_i^{(l)}, \theta_{-i}) \) without knowing the channel distributions or transmission policies of other nodes.

D. Convergence of (34) and Discussion

Define a local maximizer of the objective function (35) as

\[
\theta_i^{*} = \left\{ \theta_i : \nabla_{\theta_i} T_i(\theta_i, \theta_{-i}) = 0, \nabla_{\theta_i}^2 T_i(\theta_i, \theta_{-i}) < 0 \right\}. \tag{38}
\]

Theorem 5: For each node \( i \) and a fixed outer loop index \( n \), if \( \Delta \geq 1/\epsilon_i \), the sequence \( \theta_i^{(l,n)}, l = 1, 2, \ldots, \Delta \), generated by (34) converges in probability (i.e., weakly) to the threshold level corresponding to a local optimizer \( \theta_i^{*} \) of node \( i \)'s expected reward \( T_i(\theta_i, \theta_{-i}) \), where \( \theta_{-i} = ((\theta_{-i}^{(m)})_{j \neq i}, (\theta_{-i}^{(m-1)})_{j \neq i}) \).

Proof: For a fixed initial \( \{\theta_i^{(l,n)}\}_l \) the sequence \( \nabla_{\theta_i^{(l,n)}} T_i(\theta_i^{(l,n)}) = -\hat{\Psi}_i^{(l,n)}(\theta_i^{(l,n)}) - c_t^i + c_w^i \) for \( l = 1, 2, \ldots, \Delta \) is independent and identically distributed (i.i.d.). Therefore, (34) is an instance of the well-known constant step size Robbins Monro algorithm. Define the continuous-time interpolation of the discrete sequence \( \{\theta_i^{(l,n)}\}_l \) as follows (see [29] for details):

\[
\theta_i^{(t)}(t) = \begin{cases} \theta_i^{(l,n)}, & \text{for } t < 0 \vspace{0.1cm} \\ \theta_i^{(l,n)}, & \text{for } l \epsilon_i - \epsilon_i \leq t < l \epsilon_i, \quad l = 1, \ldots, \Delta \end{cases}. \tag{39}
\]

In [29, Theorem 2.1, p. 219], it is shown that under the condition of uniform integrability of \( \nabla_{\theta_i} T_i(\theta_i, \theta_{-i}) = \)
\[-(\Psi_i(\theta_{i}, \theta_{-i}) = c_i^{(1)} + c_i^{(2)} f_i(\theta_i)\] and four other conditions that are trivially satisfied (conditions A1.2, A1.3, A1.4, A1.10 in [29, p. 217]), as $\varepsilon_i \to 0$, $\theta_{i}^{(t)}$ defined as above converges weakly in trajectory (on the space of continuous on the right, limit on the left functions, endowed with a Skorohod metric) to the continuous-time function $\tilde{\theta}_i(t)$, i.e., $\lim_{t \to 0} \mathbb{P}(\sup_{0 \leq s \leq t} |\tilde{\theta}_i(t) - \theta_i(t)| > 0) \to 0$. $\tilde{\theta}_i(t)$ satisfies
\[
\frac{d\tilde{\theta}_i}{dt} = -\left(\Psi_i(\tilde{\theta}_i, \theta_{-i}) - c_i^{(1)} + c_i^{(2)} f_i(\theta_i). \right) \tag{40}
\]
A sufficient condition for uniform integrability of $\nabla \theta_i T_i(\theta_{i}, \theta_{-i})$ is that the channel distribution has finite variance.

The ordinary differential (40) is asymptotically stable in the sense of Liapunov due to the fact that $T_i(\theta_{i}, \theta_{-i})$ is bounded, quasi-concave with respect to $\theta_{i}$, and a local maximizer of $T_i(\theta_{i}, \theta_{-i})$ is also its global maximizer (see the remarks after the proof of Theorem 2).

Therefore, on a time scale $\Delta = 1/\varepsilon_i \to \infty$ as $\varepsilon_i \to 0$, the sequence of transmission thresholds generated by (34) converges weakly to a local maximizer of the utility function $T_i(\theta_{i}, \theta_{-i})$. In [29], the weak convergence proofs for much more general correlated signals (e.g., Markovian signals) are given.

Since at every outer loop iteration, the nodes take turns to update their policies and for each node the estimates generated by (34) converges in probability to a global optimizer of its utility function, Algorithm 2 converges weakly in trajectory to the deterministic best response dynamic adjustment process. Hence, if Algorithm 1 converges to a Nash equilibrium transmission threshold vector $\tilde{\theta}^{*}$, Algorithms 2 also converges in probability to $\tilde{\theta}^{*}$ [30].

Remark: For the convergence of Algorithm 2, we have the condition that $\Delta \geq 1/\varepsilon_i$. When $\Delta$ is small, Algorithm 2 does not exhibit the best response dynamic adjustment process for the transmission game but may still converge to a Nash equilibrium. In fact, if a small value of $\Delta$ is used, the algorithm may converge at a much faster rate and may be able to track changes in the network more efficiently. We illustrate the convergence of Algorithm 2 with $\Delta = 1$ via a numerical example in Section VII-B.

VII. NUMERICAL STUDIES

In this section, the structural results in Sections IV, V and the performance of Algorithms 1 and 2 are illustrated via numerical examples. It is assumed that the SNR represents the channel state. We simulated the time-slotted CDMA network with matched filter receivers and the SINR threshold reception model. The SINR threshold reception model is an instance of the G-MPR model and can be summarized as follows: in a transmission time slot, a packet from node $i$ is considered successfully received if and only if
\[
I \left(\frac{\gamma_i}{1 + \sum_{j \in N} \gamma_j} > \beta \right) : i, j \in A_i^K
\]
where $A_i^K$ is the set of indexes of all transmitting nodes, $\gamma_i$ is the channel state, i.e., SNR, of node $i$ and $\beta$ represents the QoS requirement.

We assume that the underlying physical channel is Rayleigh fading, hence the channel state (SNR) is exponentially distributed, i.e., the channel distribution function for node $i$ is $F(\gamma_i) = 1 - e^{-\gamma_i/\mu_i}$, where $\mu_i$ is the mean value of $\gamma_i$, $i = 1, 2, \ldots, K$. Each noncooperative game is then completely defined by 6 parameters: the number of active nodes in the network $K$, the means $\mu_i$, $i = 1, 2, \ldots, K$, of the exponential channel distributions of all nodes in the network, the spreading gain of the CDMA code sequences $N$, the QoS requirement parameter $\beta$, the transmission and waiting costs for each node.

In order to analyse the simulation results, for every numerical example, we precompute (i.e., compute offline) a Nash equilibrium profile using a modified version of Algorithm 2 with $\Delta = 1$ and a stopping criterion: the iterations are stopped when no player can gain more than $\epsilon$ by deviating from its current policy with $\epsilon = 0.0005$. In the numerical examples presented below, Algorithms 1 and 2 always lead to the same neighborhood of a Nash equilibrium. However, this does not imply the unique existence of a Nash equilibrium. In fact, studies of simple two-player transmission games will give examples of systems that have more than one Nash equilibria.

A. Existence of a Unique Nash Equilibrium in a Simple Three-Player Transmission Game

We compute the best-response functions $H(\theta_{-i}) : [0, M]^{K-1} \to [0, M]$ [defined by (28)] analytically using (31) for a simple three-player transmission game and plot the best-response functions for all three players in Fig. 2. The parameters of this game are as follows: the nodes have the same exponential channel distribution function with mean $\mu = 10$, the reception model is the SINR threshold reception model (7) for a CDMA network with matched filter receivers and the spreading gain $N = 16$ and the QoS requirement parameter $\beta = 3$ (i.e., 4.7 dB). It can be seen from Fig. 2 that $H(\theta_{-i})$ is indeed nonincreasing in every component of $\theta_{-i}$. Furthermore, Fig. 2 suggests that this specific transmission game has a unique
symmetric Nash equilibrium as the best response functions intersect at a unique point in the graph. In addition, computing the eigenvalues of the gradient matrix define in Lemma 2 at this Nash equilibrium reveals that it is asymptotic stable.

B. Convergence and Adaptation Capability of Algorithm 2 With $\Delta = 1$

The convergence and adaptive capability of Algorithm 2 with $\Delta = 1$ are illustrated in Fig. 3. The fixed parameters of the simulation are as follows: the spreading gain $N = 16$, the QoS requirement parameter $\beta = 4$ dB, the waiting and transmission costs are fixed at 0 and 0.2. Initially, the network has six active nodes that have a common exponential channel distribution function with mean $\mu = 8$. Every node uses (34) with the constant step size $\varepsilon = 0.008$ to adaptively estimate its optimal transmission threshold. The chosen initial transmission threshold is $\theta^{(0)} = 2$. Each node is programmed to send a pilot signal once every 0 transmission time slots. A pilot signal is simply a signal transmitted with exact power control. A node then updates its policy immediately based on the feedback (i.e., ACK or NACK) obtained from this pilot signal, i.e., $m = 1$. The time slots at which a node sends a pilot signal is programmed as follows: node $i$ sends a pilot signal at transmission time slot $n$ if and only if $n \mod 10 = i$, i.e., the nodes take turn to update their policies. In other words, the nodes are updating their policies according to Algorithm 2 with $m = 1$ and $\Delta = 1$.

A node is scheduled to leave the system at time slot $L = 100\,000$. A new node with the same channel distribution function as other nodes in the system (i.e., exponential distribution with mean $\mu = 8$), arrives at time slot $L = 280\,000$. Two more nodes become active at time slot $L = 500\,000$. One of these two nodes has an exponential channel distribution function with mean $\mu = 6$, and the other node has an exponential channel distribution with mean $\mu = 14$. Fig. 3 shows that every time a node becomes active or inactive, the nodes in the system update their transmission thresholds to the new equilibrium in approximately $T_1 \approx T_2 \approx T_3/3 \approx 10\,000$ transmission time slots. In real time, $10\,000$ transmission time slots are in order of seconds assuming a bandwidth of 1 MHz and a packet size of 1000 bits. It can also be seen that the two nodes that become active at time slot $L = 500\,000$ have two different transmission thresholds at the equilibrium point. This is predictable as the channel distribution functions of these two nodes differ from each other and from other nodes. In comparison, the nodes that have the same channel distribution function deploy the same transmission threshold at equilibrium.

C. Deterministic Best-Response Dynamic and Emergence of a Nash Equilibrium

We illustrate the convergence of Algorithm 1 via a numerical example. Specifically, we consider a symmetric network of five nodes with a common exponential channel distribution with mean $\mu = 8$, the spreading gain is $N = 16$, the QoS requirement $\beta = 4$ dB, the waiting and transmission costs are fixed at 0 and 0.2 respectively. The mapping (29) is numerically computed using (34). The stopping criterion is $\max_i |\theta^{(i+1)}_i - \theta^{(i)}_i| < e = 0.01$. The initial points are randomly generated. Algorithm 2 stops after 49 iterations. In Fig. 4, the estimates of the best response transmission thresholds in all 49 rounds of updating transmission policies are plotted. It can be observed that the best response transmission thresholds are quickly updated to a good neighborhood of a Nash equilibrium.

D. System Throughput Improvement

Fig. 5 illustrates the gain in the system throughput at a Nash equilibrium point, which is estimated using Algorithm 2 with $\Delta = 1$. The network model parameters are: the nodes are asymmetric and have exponential channel distributions with means $\mu = 10 - K/4 + 0.5 : 0.5 : 10 + K/4$ (i.e., $\mu_i = 10 - K/4 + i/2, i = 1, 2, \ldots, K$), where $K$ is the number of active
nodes in the network, the spreading gain is $N = 32$, the QoS requirement $\beta = 4$ dB, the waiting and transmission costs are fixed at 0 and 0.2, respectively. The empirical system throughputs at Nash equilibria are estimated for $K = 2, 3, \ldots, 40$. It is shown in Fig. 5 that the Nash equilibrium system throughput is improved with the number of active nodes. In comparison, when there is no transmission control, i.e., each node is greedy and always attempts to transmit the system throughput degrades as the system becomes more heavily loaded. In addition, when decentralized CSI is not available (or not exploited) and each node transmits with a constant probability as proposed in [3], the system throughput also slightly decreases when the network size increases to close to and above the spreading gain $N = 32$.

E. Nash Transmission Threshold Versus Number of Active Nodes and Transmission Cost

In this example, a symmetric random access network is considered. The parameters of the simulation are: the nodes are symmetric and have a common exponential channel distribution with mean $\mu = 10$, the spreading gain is $N = 32$, the QoS requirement parameter $\beta = 4$ dB, the waiting cost is fixed at 0, the transmission cost is varied from 0.1 to 0.3. The symmetric Nash transmission threshold is estimated for network sizes from 5 to 32. It can be seen from Fig. 6 that as the number of active nodes or the transmission cost increase, the symmetric Nash equilibrium transmission threshold also increases. This is intuitive since an increase in the transmission cost implies that it is more expensive to transmit and each node should transmit with a less probability. The monotonic dependency of the symmetric Nash equilibrium transmission threshold on the number of active nodes has been proved in Theorem 4.

VIII. CONCLUSION

In this paper, we exploit decentralized CSI for optimal transmission policies in multipacket reception random access networks using a game theoretic approach. The objective of each node is to select a transmission policy mapping CSI to transmission probabilities to maximize its utility. We prove the optimality of threshold transmission policies and the existence of a Nash equilibrium at which every node deploys a threshold transmission policy. For the symmetric network model we prove that there exists a Nash equilibrium profile at which all nodes deploy the same threshold transmission policy.

We study the deterministic best response dynamics algorithm for the transmission game and prove its convergence for the two-user case. We explicitly characterize the best response functions for the transmission game of an arbitrary number of nodes considering Rayleigh fading channels. This characterization offers a means to numerically verify local asymptotic stability of a Nash equilibrium. However, to a large extent, convergence of the best response dynamics algorithm has not been completely studied in this paper, and will be an interesting problem for future work. At each iteration of the best response dynamics algorithm, a node has to solve a stochastic optimization problem to update its policy to the best response transmission policy. We propose an algorithm that uses the gradient ascent method (with a constant step size, which enhances the tracking capability) to estimate the best response transmission policy for a node at every iteration. This algorithm has the capability to track slowly time varying networks.

It is possible to establish the structure of the optimal transmission policy for a more general network model with MPR and packet priorities based on the results presented in this paper. In addition, in [15], we consider the problem of centralized optimal transmission control to optimize the system throughput and study the structure of the optimal transmission policy.

APPENDIX

A. Proof of Theorem 2

When every node considers threshold transmission policies only, the space of pure policies is the set of possible values of the
transmission thresholds, i.e., $[0, M]$, which is a compact, convex subset of $\mathbb{R}_+$.

A classical result on the existence of a Nash equilibrium for games with continuous utility functions [23]–[25], also see [18, Theorem 1.2]) states that if the space of pure policies of each player is a nonempty, compact, convex subset of an Euclidean space, in order to show the existence of a Nash equilibrium, it suffices to show that the utility function for each node is continuous with respect to all policies deployed by other nodes in the network and quasi-concave with respect to its policy. Consider node $i$. Rewrite the utility function of node $i$, given by (22), as

$$
T_i(\theta_{-i}, \theta_{-i}) = \int_{\delta_i}^{M} \prod_{k=1}^{K} \left[ (1 - \mathbf{I}(\gamma_k > \theta_k)) \psi_i \left( \gamma_i, \theta_{-i} \right) \right] dF_i(\gamma_i) - c^{(i)}_d - c^{(i)}_w.
$$

(41)

It can be easily seen that the only role of $\theta_j$, $j = 1, 2, \ldots, K$ in (41) is to change the limits of the integrals. Also note that $\psi_i(\gamma_i, \theta_{-i})$ is Lebesgue measurable since the function is Lebesgue measurable, defined over a bounded domain and has a bounded range. By the Fundamental Theorem of Calculus [28], $T_i(\theta_{-i}, \theta_{-i})$ is absolutely continuous with respect to all $\theta_j$, $j = 1, 2, \ldots, K$.

We now prove that $T_i(\theta_{-i}, \theta_{-i})$ is quasi-concave with respect to $\theta_i$. Indeed, we have

$$
T_i(\theta_{-i}, \theta_{-i}) = \int_{\delta_i}^{M} \prod_{k=1}^{K} \left[ (1 - \mathbf{I}(\gamma_k > \theta_k)) \psi_i \left( \gamma_i, \theta_{-i} \right) \right] dF_i(\gamma_i) - c^{(i)}_d - c^{(i)}_w
$$

(42)

where $\psi_i(\gamma_i, \theta_{-i})$ is given by (15) with $p_j(\gamma_j) = \mathbf{I}(\gamma_j > \theta_j)$

$$
\Psi_i(\gamma_i, \theta_{-i}) = \mathbb{E}_{\mathbf{\gamma}_{-i}} \left[ \sum_{k=1}^{K} \sum_{i \in \mathcal{A}_k^i} \left( \prod_{j \neq i, j=1}^{K} (1 - \mathbf{I}(\gamma_j > \theta_j)) \psi_i \left( \gamma_i, \theta_{-i} \right) \right) \right].
$$

(43)

Due to (8), (11), and (13), it is clear that $\psi_i(\gamma_i, \theta_{-i})$ is nondecreasing in $\gamma_i$ and $\Psi_i(0, \theta_{-i}) - c^{(i)}_d + c^{(i)}_w < 0$. Consider two cases:

1) If $\Psi_i(M, \theta_{-i}) - c^{(i)}_d + c^{(i)}_w \leq 0$, then

$$
\psi_i(\theta_i, \theta_{-i}) - c^{(i)}_d - c^{(i)}_w \begin{cases} < 0, & \text{for all } \theta_i < \lambda, \\
= 0, & \text{for all } \theta_i = \lambda, \\
> 0, & \text{for all } \theta_i > \lambda
\end{cases}
$$

for some $\lambda \in (0, M]$. Consequently, $T_i(\theta_i, \theta_{-i})$ is strictly increasing for $\theta_i < \lambda$ and constant for $\theta_i > \lambda$. Overall, $T_i(\theta_i, \theta_{-i})$ is a nonempty, compact, convex subset of domain $[0, M]$. That is, $T_i(\theta_i, \theta_{-i})$ is quasi-linear and hence quasi-concave in $\theta_i$.

2) If $\Psi_i(M, \theta_{-i}) - c^{(i)}_d + c^{(i)}_w \geq 0$, then there exists a $\lambda \in (0, M]$ such that

$$
\psi_i(\theta_i, \theta_{-i}) - c^{(i)}_d - c^{(i)}_w \begin{cases} < 0, & \text{for all } \theta_i < \lambda, \\
= 0, & \text{for all } \lambda \leq \theta_i \leq \lambda_2, \\
> 0, & \text{for all } \theta_i > \lambda_2
\end{cases}
$$

$T_i(\theta_i, \theta_{-i})$ is then increasing for $\theta_i \in (0, \lambda_1)$, constant for $\theta_i \in (\lambda_1, \lambda_2)$, and decreasing for $\theta_i \in (\lambda_2, M]$. Recall that $T_i(\theta_i, \theta_{-i})$ is also continuous in $\theta_i$. Hence, $T_i(\theta_i, \theta_{-i})$ is quasi-concave in $\theta_i$.

Therefore, a Nash equilibrium exists.

**Remarks:**

1) From the properties that are listed above for $T_i(\theta_i, \theta_{-i})$, it can be seen that even though the function is only quasi-concave in $\theta_i$, any local maximizer $\theta^*_{-i}$ of $T_i(\theta_i, \theta_{-i})$ is also a global maximizer.

2) If (9) holds instead of (8) then the utility function $T_i(\theta_i, \theta_{-i})$ is strictly quasi-concave with respect to $\theta_i$.

**B. Proof of Theorem 3**

In this proof, we make use of the Kakutani’s fixed-point theorem [31], [17, Lemma 20.1].

**Lemma 4:** Let $X$ be a compact, convex subset of $\mathbb{R}^n$ and $f : X \rightarrow X$ be a set-valued function that has the following conditions:

- for all $x \in X$, the set $f(x)$ is nonempty and convex;
- the graph of $f$ is closed (i.e., for all sequences $\{x_n\}$ and $\{y_n\}$ such that $y_n \in f(x_n)$ for all $n$, $x_n \rightarrow x$, $y_n \rightarrow y$, we have $y \in f(x)$).

Then, there exists $x^* \in X$ such that $x^* \in f(x^*)$.

Using the Kakutani’s fixed-point theorem, it will be shown that there exists a $\theta^* \in [0, M]$ such that all nodes other than node $i$ in the network deploy the threshold transmission policy $p^*(\gamma) = \mathbf{I}(\gamma > \theta^*)$, then $p^*(\cdot)$ is also an optimal policy for node $i$.

Assume that nodes in the set $\{1, 2, \ldots, K\} - \{i\}$ use the transmission policy $p(\gamma) = \mathbf{I}(\gamma > \theta)$. In order to maximize its utility, node $i$ must solve the optimization problem $\max_{\theta_i \in [0, M]} T_i(\theta_i, \theta_{-i})$, where $T_i(\theta_i, \theta) = \int_{\delta_i}^{M} \mathbf{I}(\gamma > \theta)(\Psi_i(\gamma_i, \theta) - c_d + c_w) dF(\gamma_i)$.

Due to the continuity, monotonicity of $\psi_i(\gamma_i, \theta)$ and the bang-bang principle in [22], we have

$$
\arg \max_{\theta_i \in [0, M]} T_i(\theta_i, \theta_{-i}) = \{\theta_i : \Psi_i(\theta_i, \theta_{-i}) - c_d + c_w = 0\}. 
$$

(43)

Define a set-valued inverse function of $\psi_i(\cdot, \cdot)$ as follows:

$$
f(\theta) \hat{=} \Psi_i^{-1}(0, \theta) = \{\gamma : \psi_i(\gamma_i, \theta) - c_d + c_w = 0\} : [0, M] \rightarrow [0, M].
$$

(44)

Equation (43) can then be rewritten as

$$
\arg \max_{\theta_i \in [0, M]} T_i(\theta_i, \theta_{-i}) = f(\theta).
$$

We now prove that $f(\theta)$ satisfies the conditions of the Kakutani’s fixed-point theorem. Recall that $\psi_i(\gamma_i, \theta)$ is continuous and nondecreasing in $\gamma_i$. Note that the conditions (11) and (27) imply that

$$
\Psi_i(0, \theta) - c_d + c_w < 0 \quad \text{and} \quad \Psi_i(M, \theta) - c_d + c_w > 0.
$$

(45)
It follows from inequalities in (45) and the continuity of $\Psi_i(\gamma_i, \theta)$ with respect to $\gamma_i$ that $f(\theta) \neq \varnothing$ for every $\theta \in [0, M]$. In addition, due to the continuity and monotonicity of $\Psi(\gamma, \theta)$ with respect to $\gamma$, $f(\theta)$ defined as in (44) is either a singleton set containing only one real number, or a closed, bounded interval of $\mathbb{R}_+$. Hence, $f(\theta)$ is a compact, convex set for all $\theta \in [0, M]$. In addition, the closure of the graph of $f(\theta)$ follows straightforwardly from the continuity of $\Psi(\gamma, \theta)$. Therefore, by the Kakutani’s fixed-point theorem, there exists a $\theta^*$ such that $\theta^* \in f(\theta^*)$. Therefore, a symmetric Nash equilibrium exists.

C. Proof of Theorem 4

To distinguish between the two numbers of active nodes the superscripts $(K)$ and $(K + 1)$ are added to refer to the system of $K$ and $K + 1$ active nodes, respectively. From (44), it is clear that $\eta$ and $\xi$ are the Nash equilibrium transmission thresholds for the transmission game of $K$ and $K + 1$ players, respectively, if and only if $f_i(K)(\eta_i, \eta) - c_t + c_w = 0$ and $f_i(K + 1)(\xi, \xi) - c_t + c_w = 0$. Hence, is a compact, convex set for all $i$. In addition, the closure of the graph of follows straightforwardly from the continuity of $f(\theta)$. Therefore, by the Kakutani’s fixed-point theorem, there exists a $\theta^*$ such that $\theta^* \in f(\theta^*)$. Therefore, a symmetric Nash equilibrium exists.

D. Proof of Lemma 1

For a two-player game, in the $n$th round of updating policies, the deterministic best response dynamic can be written as below:

\[
\begin{align*}
\theta_i^{(n)}(\cdot) &= H\left(\theta_i^{(n-1)\ell}\right) = \gamma: \Psi_i\left(\gamma, \theta_i^{(n-1)\ell}\right) - c_t^{(n)} + c_w^{(n)} = 0 \\
\theta_i^{(n)(\cdot)} &= H\left(\theta_i^{(n)}\right) = \gamma: \Psi_i\left(\gamma, \theta_i^{(n)}\right) - c_t^{(2)} + c_w^{(2)} = 0,
\end{align*}
\]

Due to the monotonicity of the best response function $H(\cdot)$, it is easy to see that $\theta_i^{(n+1)} \geq \theta_i^{(n)} \geq \theta_i^{(n+1)} \geq \theta_i^{(n+2)} \geq \theta_i^{(n+1)}$, and $\theta_0^{(n)} \leq \theta_0^{(n)} \leq \theta_0^{(n+1)} \leq \theta_0^{(n+2)} \leq \theta_0^{(n+1)}$

Hence, after the first iteration, the best response dynamic proceeds monotonically in both directions. Since the transmission thresholds only take values in $[0, M]$, the best response dynamic converges.

E. Proof of Lemma 3

We prove (31) by mathematical induction. The proof of (32) follows the same steps but is longer and its details are omitted. Rewrite (25) as follows:

\[
\Psi_i(\gamma_i, \theta_i^{-}) = \sum_{k=1}^{K} \sum_{j \notin A_i^k} \prod_{l \neq j} (1 - e^{-\lambda_j t_l}) \times P\left( a > \gamma_j \text{ and } \gamma_j > \gamma_j^\ast j \in A_i^k \setminus \{i\} \right).
\]

In order to prove (31), we need to prove for all $n \geq 1$ that

\[
P\left( a > \sum_{j=1}^{n} \gamma_j, \gamma_1 > \gamma_1^\ast > \gamma_1^\ast n \right) = \mathcal{I} \left( a > \sum_{j=1}^{n} \theta_j \right) \exp\left( -\gamma_1^\ast \sum_{j=1}^{n} \theta_j \right) - \exp(-\gamma_1^\ast a) + \sum_{l=1}^{n-1} \frac{\lambda e^{-\gamma_1^\ast a} (a - \sum_{j=1}^{n} \theta_j)^l}{(-1)^l l!}
\]

where $\gamma_1^\ast = (\gamma_1, \ldots, \gamma_n)$, $\gamma_1^\ast n = (\theta_1, \ldots, \theta_n)$. Indeed, (46) holds for $n = 1$. Assume that (46) holds for $n$; we now prove that it holds for $n + 1$.

\[
P\left( a > \sum_{j=1}^{n+1} \gamma_j, \gamma_1 > \gamma_1^\ast > \gamma_1^\ast n \right) = \mathcal{I} \left( a > \sum_{j=1}^{n+1} \theta_j \right) A
\]

where $\gamma_1^\ast = (\gamma_1, \ldots, \gamma_{n+1})$, $\gamma_1^\ast n = (\theta_1, \ldots, \theta_{n+1})$, and

\[
A = \int_{\theta_1^{n+1}}^{\sum_{j=1}^{n+1} \theta_j} \mathcal{P}\left( a > \gamma_{n+1} + \sum_{j=1}^{n} \gamma_j, \gamma_1 > \gamma_1^\ast n \right) \times \mathcal{E}^{-\lambda_1 \gamma_1^\ast n} d\gamma_{n+1}
\]

\[
= \int_{\theta_1^{n+1}}^{\sum_{j=1}^{n+1} \theta_j} -\lambda e^{-\lambda_1 a} + \lambda e^{-\lambda_1 \gamma_1^\ast n} \left( e^{-\gamma_1^\ast \sum_{j=1}^{n} \theta_j} \right) \mathcal{E}^{-\lambda_1 \gamma_1^\ast n} d\gamma_{n+1}
\]

\[
= e^{-\gamma_1^\ast a} - e^{-\gamma_1^\ast a} + \sum_{l=1}^{n} \frac{\lambda e^{-\gamma_1^\ast a} \left( a - \sum_{j=1}^{n} \theta_j \right)^l}{(-1)^l l!}
\]

The proof of (32) follows exactly the same steps. Instead of (46), we have to prove the inequality below by mathematical induction. Due to the lack of space, the details are omitted.
REFERENCES


Minh Hau Ng¢ (S’04) was born in 1979. She received the B.E. degree in telecommunications engineering from the University of Sydney, Australia, in 2003. She is currently working towards the Ph.D. degree at the University of British Columbia, Vancouver, BC, Canada.

Her current research interests include stochastic optimization, game theory, and wireless communications.

Ms. Ngo received a Killam Pre-doctoral Fellowship from the University of British Columbia.

Vikram Krishnamurthy (S’80–M’91–SM’99–F’05) was born in 1966. He received the Bachelor’s degree from the University of Auckland, New Zealand, in 1988 and the Ph.D. degree from the Australian National University, Canberra, in 1992.

Previously, he was a chaired Professor at the Department of Electrical and Electronic Engineering, University of Melbourne, Australia. Since 2002, he has been a Professor and Canada Research Chair at the Department of Electrical Engineering, University of British Columbia, Vancouver, BC, Canada. His research interests span several areas, including ion channels and nanobiology, stochastic scheduling and control, statistical signal processing, and wireless telecommunications.

Krishnamurthy has served as Associate Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS AEROSPACE AND ELECTRONIC SYSTEMS, IEEE TRANSACTIONS ON NONOBIOCIENCE, IEEE TRANSACTIONS CIRCUITS AND SYSTEMS II, Systems and Control Letters, and the European Journal of Applied Signal Processing. He was Guest Editor of the Special Issue on Bionanotubes of IEEE TRANSACTIONS ON NONOBIOCIENCE (March 2005).