Competitive scheduling in wireless collision channels with correlated channel state

Citation

As Published
http://dx.doi.org/10.1109/GAMENETS.2009.5137452

Publisher
Institute of Electrical and Electronics Engineers

Version
Final published version

Accessed
Sun Apr 14 09:24:07 EDT 2019

Citable Link
http://hdl.handle.net/1721.1/58939

Terms of Use
Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use.

Detailed Terms

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Competitive Scheduling in Wireless Collision Channels with Correlated Channel State

Utku Ozan Candogan, Ishai Menache, Asuman Ozdaglar and Pablo A. Parrilo

Abstract—We consider a wireless collision channel, shared by a finite number of mobile users who transmit to a common base station. Each user wishes to optimize its individual network utility that incorporates a natural tradeoff between throughput and power. The channel quality of every user is affected by global and time-varying conditions at the base station, which are manifested to all users in the form of a common channel state. Assuming that all users employ stationary, state-dependent transmission strategies, we investigate the properties of the Nash equilibrium of the resulting game between users. While the equilibrium performance can be arbitrarily bad (in terms of aggregate utility), we bound the efficiency loss at the best equilibrium as a function of a technology-related parameter. Under further assumptions, we show that sequential best-response dynamics converge to an equilibrium point in finite time, and discuss how to exploit this property for better network usage.

I. INTRODUCTION

Wireless technologies are broadly used today for both data and voice communications. The transmission protocols of wireless devices need to cope with the scarce resources available, such as bandwidth and energy. Additional difficulties relate to the dynamic nature of wireless networks. For example, the mobility of terminals and the frequent change in their population introduce new challenges for routing protocols. An additional distinctive dynamic feature of wireless communications is the possible time variation in the channel quality between the sender and the receiver, an effect known as channel fading [4].

Motivated by scalability considerations, it has been advocated in recent years that mobiles should have the freedom to distributively adjust their transmission parameters (e.g., [8]). This leads to distributed network domains, in which situation-aware users take autonomous decisions with regard to their network usage, based on the current network conditions and their individual preferences. As users generally do not coordinate their network actions, non-cooperative game-theory has become a natural theoretical framework for analysis and management (see, e.g., [2], [1], [6], and [8] for a survey). The general framework that will be considered here is that of users who obtain state information and may accordingly control their transmission parameters. A major issue in this context is whether such information is well exploited by self-interested users, or rather misused.

In this paper, we consider a wireless collision channel, shared by a finite number of users who wish to optimally schedule their transmissions based on a natural tradeoff between throughput and power. The channel quality between each user and the base station is randomly time-varying and observed by the user prior to each transmission decision. The bulk of the research in the area has been carried under a simplified assumption that the channel state processes of different users are independent (see e.g., [3], [7]). In practice, however, there are global system effects, which simultaneously affect the quality of all transmissions (e.g., thermal noise at the base station, or common weather conditions). Consequently, a distinctive feature of our model is that the state processes of different users are correlated. As an approximating model, we consider in this paper the case of full correlation, meaning that all users observe the same state prior to transmission. A fully correlated state can have a positive role of a coordinating signal, in the sense that different states can be “divided” between different users. On the other hand, such state correlation increases the potential deterioration in system performance due to non-cooperation, as users might transmit simultaneously when good channel conditions are available.

Our results indicate that both the positive and negative effects of state correlation are possible. User interaction and unwillingness to give up the better-quality states leads to arbitrarily bad equilibria in terms of the aggregate utility. At the same time, there exist good-quality equilibria, whose performance-gap from the social welfare solution can be bounded as a function of a technology-related parameter. In some special cases, we establish that the underlying scheduling game is a potential game [10]. We use this property and additional structure of our game, to establish the finite-time convergence of sequential best-response dynamics (although the action space is continuous). We further show by simulations that convergence to an equilibrium can be very fast. We use the latter property to indicate how good-quality equilibria can be iteratively obtained.

The structure of the paper is as follows. The model and some basic properties thereof are presented in Section II. The social welfare problem is defined and characterized in Section III. We then proceed in Section IV to study the efficiency loss under selfish user behavior, as compared to the social welfare operating point. Section V studies the convergence properties of best-response dynamics and their significance. Conclusions are drawn in Section VI.

II. THE MODEL AND PRELIMINARIES

We consider a wireless network, shared by a finite set of mobile users \( M = \{1, \ldots, M\} \) who transmit at a
fixed power level to a common base station over a shared collision channel. Time is slotted, so that each transmission attempt takes place within slot boundaries that are common to all. Transmission of a user is successful only if no other user attempts transmission simultaneously. Thus, at each time slot, at most one user can successfully transmit to the base station. To further specify our model, we start with a description of the channel between each user and the base station (Section II-A), ignoring the possibility of collisions. In Section II-B, we formalize the user objective and formulate the non-cooperative game which arises in a multi-user shared network.

A. The Physical Network Model

Our model for the channel between each mobile (or user) and the base station is characterized by two basic elements.

a. Channel state process. We assume that the channel state between mobile $m$ and the base station evolves as a stationary process $H^m(t), t \in \mathbb{Z}_+$ (e.g., Markovian) taking values in a set $\mathcal{H}^m = \{1, 2, \ldots, h^m\}$ of $h^m$ states. The stationary probability that mobile $m$ would observe state $i \in \mathcal{H}^m$ at any time $t$ is given by $\pi_i^m$.

b. Expected data rate. We denote by $R_i^m > 0$ the expected data rate (or simply, the rate)$^1$ that user $m$ can sustain at any given slot as a function of the current state $i \in \mathcal{H}^m$. We further denote by $\mathcal{R}^m = \{R_1^m, R_2^m, \ldots, R_{h^m}^m\}$ the set of all data rates for user $m$, and define $\mathcal{R} = \mathcal{R}^1 \times \mathcal{R}^2 \times \cdots \times \mathcal{R}^M$. For convenience, we assume that for every $m \in \mathcal{M}$ the expected data rate $R_i^m$ strictly decreases in the state index $i$, so that $R_1^m > R_2^m > \cdots > R_{h^m}^m$, i.e., state 1 represents the “best state” in which the highest rate can be achieved.

Note that the actual channel quality may still take continuous values, which each user reasonably quantifies into a finite number of information states. Using the cumulative density function of the underlying channel quality, the expected data rates and their associated steady state probabilities can be obtained. The motivation behind considering a discrete state process rather than the actual channel quality is the technical infeasibility of mobiles to sense and process a continuum of states.

A central assumption in this paper is that the state processes of different users are fully correlated, as we formalize below.

Assumption 1 (Full Correlation): All users observe the same channel state $H(t)$ in any given time $t$. That is, for every mobile $m \in \mathcal{M}$: (i) $\mathcal{H}^m = \mathcal{H} = \{1, 2, \ldots, h\}$, (ii) $\pi_i^m = \pi_i$ for every $i \in \mathcal{H}$ and (iii) $H^m(t) = H(t)$ (where $\mathcal{H}$ is the common state space, and $\pi = (\pi_1, \ldots, \pi_h)$ is its stationary distribution).

We emphasize that although all mobiles observe the same state, the corresponding rates $R_i^m$ need not be equal across mobiles, i.e., in our general model we do not assume that $R_i^m = R_j^m$, $m, k \in \mathcal{M}, i \in \mathcal{H}$. The case where the latter condition does hold will be considered as a special case in Section V.

The above model can be used to capture the following network scenario. The channel state corresponds to global conditions that affect all user transmissions. Examples may include thermal noise at the base station and weather conditions (that play a central role, e.g., in satellite networks), which affect all mobiles in a similar manner. The state information can be communicated from the base-station to all mobiles via a feedback channel. After obtaining the state information at the beginning of each slot, a user may respond by adjusting its coding scheme in order to maximize its data rate on that slot. The rate $R_i^m$ thus takes into account the quality of the current state $i$, the coding scheme adapted by the user, and “local” characteristics, such as the user’s transmission power, the location relative to the base station and (local) fast-fading effects. We emphasize that $R_i^m$ is an average quantity, which averages possible fluctuations in local channel conditions, which usually occur at a faster time-scale relative to the change in the global channel state (see, e.g., [9]). This assumption is commensurate with practical considerations, as mobiles usually cannot react to fast local changes.

B. User Objective and Game Formulation

In this subsection we describe the user objective and the non-cooperative game which arises as a consequence of the user interaction over the collision channel. In addition, we provide some basic properties and examples for the Nash equilibrium of the underlying game.

1) Basic Definitions: The basic assumption of our model is that users always have packets to send, yet they are free to determine their own transmission schedule in order to fulfill their objectives. Furthermore, users are unable to coordinate their transmission decisions.

Our focus in this paper is on stationary transmission strategies, in which the decision whether to transmit or not can depend (only) on the current state. A formal definition is provided below.

Definition 1 (Stationary Strategies): A stationary strategy for user $m$ is a mapping $\sigma^m : \mathcal{H} \to [0, 1]^h$. Equivalently, $\sigma^m$ is represented by an $h$-dimensional vector $p^m = (p_1^m, \ldots, p_h^m) \in [0, 1]^h$, where the $i$-th entry corresponds to user $m$’s transmission probability when the observed state is $i$.

We use the term multi-strategy when referring to a collection of user strategies, and denote by the vector $p = \{p^1, \ldots, p^M\}$ the multi-strategy comprised of all users’ strategies. The multi-strategy representing strategies of all users but the $m$th one is denoted by $p^{-m}$. For a given multi-strategy $p$, we define below the Quality of Service (QoS) measures that determine user performance. Let $W^m$ be the (fixed) transmission power of user $m$ per transmission attempt, and denote by $F^m(p^m)$ its average power investment, as determined by its strategy $p^m$. Then clearly, $F^m(p^m) = W^m \sum_{i=1}^{h} \pi_i R_i^m$ for every user $m$. We normalize the latter measure by dividing it by $W^m$, and consider henceforth the
normalized power investment, given by
\[ P^m(p^m) = \sum_{i=1}^{h} \pi_i p^m_i. \] (1)

For simplicity, we shall refer to \( P^m(p^m) \) as the power investment of user \( m \). We assume that each user \( m \) is subject to an individual power constraint \( 0 < P^m \leq 1 \), so that any user strategy \( p^m \) should obey
\[ P^m(p^m) \leq \bar{P}^m. \] (2)

The vector of power constraints is denoted by \( \bar{P} = (\bar{P}^1, \ldots, \bar{P}^M) \).

The second measure of interest is the mobile’s average throughput, denoted by \( T^m(p^m, p^{-m}) \). The average throughput of every user \( m \) depends on the transmission success probability at any given state \( i \), \( \prod_{k \neq m}(1 - p^k_i) \).

Hence,
\[ T^m(p^m, p^{-m}) = \sum_{i=1}^{h} \pi_i r^m_i p^m_i \prod_{k \neq m}(1 - p^k_i). \] (3)

Each user wishes to optimize a natural tradeoff between throughput and power, which is captured by maximizing the following utility function
\[ u^m(p^m, p^{-m}) = T^m(p^m, p^{-m}) - \lambda^m P^m(p^m), \] (4)

subject to the power constraint (2), where \( \lambda^m \geq 0 \) is a user-dependent tradeoff coefficient. We use the notation \( \lambda = (\lambda^1, \ldots, \lambda^M) \) for the vector of all users’ tradeoff coefficients; note that each game instance can now be formally described by the tuple \( I = \{M, \mathcal{R}, \pi, \lambda, \bar{P}\} \).

The term \( \lambda^m P^m(p^m, p^{-m}) \) in (4) can be viewed as the power cost of the mobile. The user utility thus incorporates both a “hard” constraint on power consumption (in the form of (2)), but also accounts for mobile devices that do not consume their power abilities to the maximum extent, as energy might be a scarce resource, the usage of which needs to be evaluated against the throughput benefit. We note that the utility (4) accommodates the following special cases:

- **Fully “elastic” users.** By setting \( P^m(p^m) = 1 \), a user practically does not have a hard constraint on power usage. Accordingly, the optimal operating point of the user is determined solely by the tradeoff between power and throughput, as manifested by the factor \( \lambda^m \).
  - The fully elastic user case has been considered in the wireless games literature in different contexts (see, e.g., [1]).
- **Power-cost neutral users.** Consider a user with \( \lambda^m = 0 \).
  - Such user is interested only in maximizing its throughput subject to a power constraint. This form of utility has been examined, e.g., in [3] and [6].

2) *Nash Equilibrium:* A strategy \( p^m \) of user \( m \) is feasible if it obeys the power constraint (2). We denote by \( E^m \) the feasible strategy space of user \( m \),
\[ E^m = \{p^m | P^m(p^m, p^{-m}) \leq \bar{P}^m, 0 \leq p^m \leq 1\}; \] (5)

the joint feasible action space is denoted by \( E = \prod_{m \in \mathcal{M}} E^m \). A Nash equilibrium point (NE) for our model is a feasible multi-strategy from which no user can unilaterally deviate and improve its utility. Formally, A multi-strategy \( p^m \) is a Nash equilibrium point if
\[ p^m \in \arg\max_{p^m \in E^m} u^m(p^m, p^{-m}), \text{ for every } m \in \mathcal{M} \] (6)

The existence a Nash equilibrium point is guaranteed, as we summarize below.

**Theorem 1:** There always exists a pure Nash equilibrium for the game.

Noting that the utility function \( u^m \), (defined in (4)) is linear in the strategy of user \( m \) and that the joint feasible region is convex, the existence result follows from a standard use of the Kakutani fixed point theorem (see, e.g., [11]).

We conclude this section by a couple of examples which point to some interesting features of the underlying game. The first example shows that there are possibly infinitely many Nash equilibria.

**Example 1:** Consider a game with two users, \( m, k \), and two states 1, 2. Let \( \pi_1 = \pi_2 = \frac{1}{2} \), \( R_1^m = R_2^m = 10 \), \( R_2^m = R_k = 5 \), \( \lambda^m = \lambda^k = 2 \), and \( \bar{P}^m = 0.8 \), \( \bar{P}^k = 0.3 \). It can be easily shown that the multi-strategy \( (p_1^m, p_2^m, p_1^k, p_2^k) = (1, 0.6, 0, x) \) is an equilibrium of the game, for any \( x \in [0.06] \).

The next example demonstrates that the behavior of the system in an equilibrium can sometimes be counterintuitive. For example, states which lead to lower expected rates can be utilized (in terms of the total power investment) more than higher quality states.

**Example 2:** Consider a game with two users, \( m, k \), and two states 1, 2. Let \( \pi_1 = \pi_2 = \frac{1}{4} \), \( R_1^m = R_1^k = 8 \), \( R_2^m = R_2^k = 3 \), \( \lambda^m = \lambda^k = 1 \), and \( \bar{P}^m = 0.8 \), \( \bar{P}^k = 0.3 \). The unique equilibrium of this game instance is given by \( (p_1^m, p_2^m, p_1^k, p_2^k) = (1, 0.6, 0, 0.6) \). Observe that the total power investment at state 1 (0.5) is lower than the total power investment at state 2 (0.6).

Both examples demonstrate some negative indications as to the predictability of the Nash equilibrium. Not only the number of equilibria is unbounded, but also we cannot rely on monotonicity results (such as total power investment increasing with the quality of the state) in order to provide a rough characterization of an equilibrium. At the same time, these observations motivate the study of performance-loss bounds at any equilibrium point, and also of network dynamics that can converge to a predictable equilibrium point. Both directions would be examined in the sequel.

### III. Social Welfare and Threshold Strategies

In this section we characterize the optimal operating point of the network. This characterization allows us to study the efficiency loss due to self-interested behavior (Section IV). The proofs for the results in this section can be found in Appendix I.

An optimal multi-strategy in our system is a feasible multi-strategy that maximizes the aggregate user utility. Formally,
\( \mathbf{p}^* \) is an optimal multi-strategy if it is a solution to the Social Welfare Problem (SWP), given by

\[
\text{(SWP)} \quad \max_{\mathbf{p} \in \mathcal{E}} u(\mathbf{p}),
\]

where

\[
u(\mathbf{p}) = \sum_m T^m(\mathbf{p}) - \lambda^m P^m(\mathbf{p}).
\]  

We note that (SWP) is a non-convex optimization problem (see [5] for a formal proof of this property). Our first result provides a per-state utilization bound for any optimal solution of (SWP).

**Proposition 1:** Let \( \mathbf{p} \) be an optimal solution of (SWP). Then \( \sum_m p_i^m \leq 1 \) for every \( i \in \mathcal{H} \).

The significance of the above result is that in case that all mobiles use the same power level \( W \) for transmission, then the total energy investment is bounded above by \( W h \), where \( h \) is the number of states. Note that this bound does not depend on the number of mobiles. The per-state utilization bound will play a key role in Section IV, while bounding the overall efficiency loss in the system.

For a further characterization of (SWP), we require the next two definitions.

**Definition 2 (Partially and Fully Utilized States):** Let \( p^m \) be some strategy of user \( m \). Under that strategy, state \( i \) is partially utilized by user \( m \) if \( p_i^m \in (0, 1) \); state \( i \) is fully utilized by the user if \( p_i^m = 1 \).

**Definition 3 (Threshold Strategies):** A strategy \( \mathbf{p}^m \) of user \( m \) is a threshold strategy if the following conditions hold: (i) User \( m \) partially utilizes at most one state, and (ii) If user \( m \) partially utilizes exactly one state, then the power constraint (2) is active (i.e., met with equality). A multi-strategy \( \mathbf{p} = (\mathbf{p}^1, \ldots, \mathbf{p}^M) \) is a threshold multi-strategy if \( \mathbf{p}^m \) is a threshold strategy for every \( m \in \mathcal{M} \).

The main result of this section is summarized below.

**Theorem 2:** There exists an optimal solution of (SWP) where all users employ threshold strategies.

Due to the non-convexity of (SWP), we cannot rely on first order optimality conditions for the characterization of the optimal solution. Nonetheless, Theorem 2 indicates that there always exist an optimal solution with some well-defined structure, which would be used in the next section for comparing the obtained performance at optimum, to an equilibrium performance.

### IV. Efficiency Loss

We proceed to examine the extent to which selfish behavior affects system performance. That is, we are interested in comparing the quality of the obtained equilibrium points to the centralized, system-optimal solution (7). Recently, there has been much work in quantifying the efficiency loss incurred by the selfish behavior of users in networked systems (see [12] for a comprehensive review). The two concepts which are most commonly used in this context are the price of anarchy (PoA), which is an upper bound on the performance ratio between the global optimum and the worst Nash equilibrium, and price of stability (PoS), which is an upper bound on the performance ratio between the global optimum and the best Nash equilibrium. The performance measure that we consider here in order to evaluate the quality of a network working point is naturally the aggregate user utility (8).

The standard definitions of PoA and PoS consider all possible instances of the associated game. Recall that in our specific framework, a game instance is given by the tuple \( I = \{ \mathcal{M}, \mathcal{R}, \pi, \lambda, \tilde{\mathbf{P}} \} \). The next example shows that the performance at the best Nash equilibrium can be arbitrarily bad compared to the socially optimal working point.

**Example 3:** Consider a network with two users \( m \) and \( k \) and two channel states. Let \( \pi_1 = \pi_2 = \frac{1}{2}, P^m = \frac{1}{4}, P^k = \frac{1}{2} \). Assume that \( R^m = R^k = \epsilon, R^m = 4, R^k = 4\epsilon, \lambda^m = \lambda^k = \frac{\epsilon}{2} \). The socially optimal working point is given by \( \mathbf{p} = (\tilde{p}_1^m, \tilde{p}_2^m, \tilde{p}_1^k, \tilde{p}_2^k) = (\frac{1}{2}, 0, 0, 1) \) and the unique equilibrium is \( \mathbf{p} = (p_1^m, p_2^m, p_1^k, p_2^k) = (0, \frac{1}{2}, 1, 0) \). Note that \( u(\tilde{\mathbf{p}}) = 1 + \frac{\epsilon}{2}, \) while \( u(\mathbf{p}) = \frac{3\epsilon}{2} \). Hence, \( \frac{u(\mathbf{p})}{u(\tilde{\mathbf{p}})} > \frac{3}{2} \), which goes to infinity as \( \epsilon \to 0 \).

The significance of the above example is that if we consider all possible game instances \( \{ \mathcal{M}, \mathcal{R}, \pi, \lambda, \tilde{\mathbf{P}} \} \), then equilibrium performance can be arbitrarily bad. However, we note that for a given mobile technology, some elements within any game instance cannot obtain all possible values. Specifically, \( \pi \) is determined by the technological ability of the mobiles to quantize the actual channel quality into a finite number of “information states” as described in section II. Naturally, one may think of several measures for quantifying the quality of a given quantization. We represent the quantization quality by a single parameter \( \pi_{max} = \min_{\mathcal{H}} \pi_i, \) with the understanding that smaller \( \pi_{max} \), the better the quantization procedure. In addition, a specific wireless technology is obviously characterized by the power constraint \( P^m \). Again, we represent the power-capability of a given technology by a single parameter \( P_{min} = \min_{m \in \mathcal{M}} P^m \). Finally, we determine the technological quality of a set of mobiles through the scalar \( Q = \frac{\pi_{max}}{P_{min}} \).

We consider next the efficiency loss for a given technological quality \( Q \). Denote by \( I_{Q_0} \) the subset of all game instances such that \( Q = Q_0 \). We define below the price of stability (PoS) and price of anarchy (PoA) as a function of \( Q \).

**Definition 4 (Price of Stability - Price of Anarchy):** For every game instance \( I \), denote by \( N_I \) the set of Nash equilibria, and let \( p^*_I \) be an optimal multi-strategy. Then for any fixed \( Q \), the PoS and PoA are defined as

\[ \text{PoS}(Q) = \sup_{I \in I_{Q_0}} \inf_{\mathbf{p} \in N_I} \frac{u(p^*_I)}{u(\mathbf{p})}, \]

\[ \text{PoA}(Q) = \sup_{I \in I_{Q_0}} \inf_{\mathbf{p} \in N_I} \frac{u(p^*_I)}{u(\mathbf{p})}. \]  

We next provide upper and lower bounds for \( \text{PoS}(Q) \) under the assumption that \( Q < 1 \) (note that an the unbounded price of stability in Example 3 was obtained for \( Q > 1 \)). The upper bound on the price of stability follows from the next proposition.
Proposition 2: Fix $Q < 1$. Let $\bar{p}$ be some threshold multi-strategy, and let $u(\bar{p})$ be the respective aggregate utility (8). Then there exists an equilibrium point $\bar{p}$ whose aggregate utility is not worse than $u(\bar{p})(1 - Q)^2$. That is, $\frac{u(p)}{u(\bar{p})} \leq (1 - Q)^2$.

The key idea behind the proof is to start from a threshold multi-strategy $\bar{p}$ and to reach an equilibrium point by some iterative process. In each step of the process we obtain the worst-case performance loss, which leads to the overall loss in the entire procedure. A full proof of this result, as well as other results in this section, can be found in Appendix II.

Recalling that there always exists an optimal threshold multi-strategy (Theorem 2), immediately establishes the following.

Corollary 1: Let $Q < 1$. Then $\text{PoS}(Q) \leq (1 - Q)^{-2}$. The above result implies that for $P_{\text{min}}$ fixed, a finer quantization of the channel quality results in a better upper bound for the PoS, which approaches 1 as $\pi_{\text{max}} \to 0$.

It is also possible to obtain a lower bound on the PoS for any given $Q$ as the next proposition suggests.

Proposition 3: Let $Q < 1$. Then $\text{PoS}(Q) \geq (1 - \frac{1}{Q+1})^{-1}$.

Observe that, for $Q \ll 1$ or for $Q = \frac{1}{\alpha} + \epsilon$ for some integer $n$ and $0 < \epsilon \ll 1$, $\frac{1}{Q+1} \approx \frac{1}{Q}$ and hence $\text{PoS} \geq \frac{1}{1-\alpha}$ for such $Q$. We can prove that $\text{PoS}(Q) \leq (1 - Q)^{-2}$ by Corollary 1, the gap between the upper and lower bound remains a subject for on-going work.

We conclude this section by showing that the PoA is unbounded for any $Q$.

Proposition 4: For any given $Q$, $\text{PoA}(Q) = \infty$.

The above result indicates that despite technological enhancements (which result in a low $Q$), the network cannot still arrive at bad-quality equilibria with unbounded performance loss. This negative result emphasizes the significance of mechanisms or distributed algorithms, which preclude such equilibria. We address this important design issues in the next section.

V. BEST-RESPONSE DYNAMICS

A Nash equilibrium point for our system represents a strategically stable working point, from which no user has incentive to deviate unilaterally. In this section we address the question of if and how the system arrives at an equilibrium, which is of great importance from the system point of view. As discussed in Section IV, the set of equilibria can vary with respect to performance. Hence, we conclude this section by briefly discussing how to lead the system to good equilibrium.

A. Convergence Properties

The most natural mechanism for (distributed) convergence to an equilibrium relies on a user’s best response (BR), which is a user’s strategy that maximizes its own utility, given the strategies of other users. In our game, a best response $\bar{p}^m$ of user $m$ to a given multi-strategy $p^m$ is given by

$$\bar{p}^m \in BR^m(p^m) = \arg\max_{\bar{p}^m \in E^m} u^m(\bar{p}^m, p^m). \quad (11)$$

An informal description of a general best-response mechanism is simple: Each user updates its strategy from time to time through a best response, (11).

The best-response mechanism, described above in its most general form, is not guaranteed to converge to an equilibrium in our system without imposing additional assumptions. We specify below the required assumptions. Our convergence analysis relies on establishing the existence of a potential function (see (10)) under a certain condition, which we refer to as the rate alignment condition. The rate alignment condition is defined as follows.

Assumption 2 (Rate Alignment Condition): The set of user rates $\{R^m_i\}_{i \in H, m \in M}$ is said to be aligned if there exist per-user positive coefficients $\{c^m\}_{m \in M}$ and per-state positive constants $\{R^m_i\}_{i \in H}$ such that

$$R^m_i = c^m R^m_i \quad (12)$$

for every $m \in M$ and $i \in H$. The rate alignment condition is satisfied if user rates are aligned.

The coefficient $c^m$ above reflects user $m$’s relative quality of transmissions, which is affected mainly by its transmission power and location relative to the base station. While the rate alignment condition might not hold for general and heterogeneous mobile systems, a special case of interest which satisfies (12) is the symmetric-rate case, i.e., $c^m = c$ for every $m \in M$. Rate-symmetry is expected in systems where mobiles use the same technology (transmission power and coding scheme), and where “local” conditions, such as distance from the base station, are similar.

For convenience below we give a definition of the ordinal potential games.

Definition 5 (Ordinal Potential Game): A game is called an ordinal potential game if a function $\phi : E \to \mathbb{R}$ such that for any $p^m, q^m \in E^m$

$$\phi(p^m, p^m) - \phi(q^m, p^m) > 0 \iff u^m(p^m, p^m) - u^m(q^m, p^m) > 0 \quad (13)$$

exists, for all $m$, and for all $p^m, q^m, p^m \in \prod_{k \neq m} E_k$. Moreover, the function $\phi$ is called an ordinal potential function.

If (13) is replaced with $\phi(p^m, p^m) - \phi(q^m, p^m) = u^m(p^m, p^m) - u^m(q^m, p^m)$ then the game is an exact potential game.

Theorem 3: Under Assumption 2, our game is an ordinal potential game with a potential function given by

$$\phi(p) = -\sum_{i=1}^{h} \pi_i R^m_i \prod_{k \in M} (1 - p^k_i) - \sum_{i=1}^{h} \sum_{k \in M} \pi_i \lambda^k p^k_i \quad (14)$$

Proof: Consider two different multi-strategies $p, q$ such that $p = (p^m, p^m), q = (q^m, p^m)$. Observe that

$$\phi(p) - \phi(q) = \frac{1}{c^m} (u^m(p) - u^m(q)) \quad (15)$$

Since $c^m > 0$, the above equality implies that the game is an ordinal potential game.
The above class of ordinal potential games are also known as weighted potential games [10]. The theorem also implies that if \( c^m = 1 \) for every \( m \in M \), then the game is an exact potential game.

In the following, we assume that users restrict themselves to threshold policies (see definition 3). Since our focus is on best response dynamics this assumption is natural as whenever a user updates its policy there always exists a threshold policy that maximizes the performance of that user. Moreover, it turns out that despite the fact that the game we are interested in is an infinite game, convergence takes place in finitely many update periods if users only use threshold policies.

Throughout this section, we assume that users may update their strategy at a slower time-scale compared to their transmission rates. For simplicity, we assume that user updates may take place only every \( T_E \) time slots and refer to \( T_E \) as the update period.

For our convergence result, we require the following set of assumptions.

**Assumption 3:**
(i) The user population is fixed.
(ii) Rates are aligned (see Assumption 2).
(iii) The transmission-success probabilities \( \prod_{k \neq i}(1 - p_k^i) \), \( i \in H \) are perfectly estimated by each user before each update.

Consider the following mechanism.

**Definition 6 (Round-Robin BR Dynamics):** A strategy update takes place in a round-robin manner and at each update period only a single user may modify its strategy. The user who is chosen for update modifies its strategy to a threshold strategy from the set \( BR^m(p^{−m}) \), if the modification strictly improves its utility.

It can be easily verified that there always exists a threshold strategy in \( BR^m(p^{−m}) \).

The next lemma suggests that our game is a finite game if users are restricted to playing threshold strategies, and further provides a bound on the number of threshold multi-strategies for any given game instance.

**Lemma 1:** For a given game instance with \( M \) users and \( h \) states the number of threshold multi-strategies is bounded by \((2e)^{M(h+1)}\).

**Proof:** See Appendix III.

Relying on the above lemma, we have the following convergence result.

**Theorem 4:** Let Assumption 3 hold. Then best response dynamics converge in finitely many update periods to an equilibrium point. In addition, the number of update periods required for convergence is upper bounded by \( M(2e)^{M(h+1)}\).

**Proof:** See Appendix III.

We emphasize that the restriction to threshold strategies is commensurate with the users’ best interest. Not only there always exists such best-response strategy, but also it is reasonably easier to implement.

We discuss next some important considerations regarding the presented mechanism and the assumptions required for its convergence. The best response dynamics as described in Definition 6, requires synchronization between the mobiles, which can be done centrally by the base station or by a supplementary distributed procedure. We emphasize that the frame of updates is the only item that needs to be centrally determined. Users are free to choose their strategies according to their own preferences, which are usually private information. Assumption 3(iii) entails the notion of a quasi-static system, in which each user responds to the steady state reached after preceding user update. This approximates a natural scenario where users update their transmission probabilities at much slower time-scales than their respective transmission rates. An implicit assumption here is that the update-period \( T_E \) is chosen large enough to allow for accurate estimation of the transmission-success probabilities. We leave the exact determination of \( T_E \) for future work. We emphasize that users need not be aware of the specific transmission probabilities \( p_k^m \) of other users. Indeed, in view of (4), only the transmission-success probabilities \( \prod_{k \neq i}(1 - p_k^i) \), \( i \in H \) are required. These can be estimated by sensing the channel and keeping track of idle slots.

A last comment relates to the rate-alignment condition. The convergence results in this section rely on establishing a potential function for the underlying game, which is shown to exist when rates are aligned. In [5], we show that in a system of three states or more, the alignment condition is not only sufficient, but also necessary for the existence of a potential. This suggests that novel methods would have to be employed for establishing convergence of dynamics under more general assumptions.

**B. Experiments**

The objective in this subsection is to study through simulations the convergence properties of sequential best-response dynamics. More specifically, we wish to examine the dependence of convergence time on several factors, such as the number of users in the system, the number of states, and the technology factor \( Q \). In all our experiments, we consider a relaxed version of Assumption 3, where the rate-alignment condition (Assumption 3(ii)) is not enforced.

The specific setup for our simulations is as follows. We assume that \( \pi_i = \frac{1}{h} \) for every \( i \in H \). For given \( Q, M \) and \( h \), we construct a significant number of game instances (10000) by randomly choosing in each instance the power constraints \( P^m \), the tradeoff coefficient \( \lambda^m \) and the associated rates \( R_i^m \) for every \( m \in M, i \in H \). We simulate each game instance, and examine the average convergence speed, measured in the number of round-robin iterations (recall in a round-robin iteration, each user updates its strategy at most once). Figure 1 presents the convergence speed results for \( Q = 0.5 \), as a function of the number of users in the system. For the given value of \( Q \), we consider three different number of states \( h \).

As seen in Figure 1, the average number of iterations required for convergence is less than three on average. We emphasize that all game instances converge without requiring the rate-alignment condition, indicating the possibility to exclude this condition in future analysis of best-response convergence. We observe that all graph curves initially
increase as a function of the number of users, and at some point gradually decrease until reaching a fixed number of iterations. This interesting phenomenon can be intuitively justified as follows: When the number of users is relatively small, there is less competition on each state, and convergence is fast. At the other extreme, when the number of users is larger than some threshold, then there are more users who can fully utilize states at the first iteration (see Definition 2), thereby decreasing the competition at subsequent iterations and leading to faster convergence.

C. Obtaining Desirable Equilibria

We conclude this section by briefly discussing possible means for obtaining high-quality equilibria in terms of the aggregate utility (8). Theorem 4 introduces a scheme (or mechanism) which assures converge to an equilibrium point in a finite number of steps. However, the resulting equilibrium can be of low quality. Proposition 2 suggests that if the system is initiated at some threshold multi-strategy, then the equilibrium performance cannot deviate by much, compared to the performance at the initial working point. Consequently, one may consider an iterative hybrid algorithm, in which a network-management entity forces some initial working point (a threshold multi-strategy), waits enough time until convergence, and if the equilibrium performance is unsatisfactory, enforces a different working point, until reaching a satisfactory equilibrium. The algorithm would rely on the fast convergence to an equilibrium, which is demonstrated in all our simulations, and allows to consider numerous initial working points in plausible time-intervals. The precise requirements and properties of such an algorithm, as well as the means for choosing and enforcing initial working-points, remain as a challenging future direction.

VI. Conclusion

We have considered in this paper a wireless network game, where mobiles interact over a shared collision channel. The novelty in our model is the state correlation assumption, which incorporates the effect of global time-varying conditions on performance. In general, the correlated state can be exploited by the users for time-division of their transmission, which would obviously increase the system capacity. However, we have shown that under self-interested user behavior, the equilibrium performance can be arbitrarily bad. Nevertheless, the efficiency loss at the best equilibrium can be bounded as a function of a technology parameter, which accounts both for the mobiles power limitations and the level of discretization of the underlying channel quality. Importantly, we have shown that best-response dynamics may converge to an equilibrium in finite time, and empirically verified that such dynamics converge fairly fast.

We briefly note several extensions and open directions: The convergence analysis of best-response dynamics under more general conditions is of great interest. As mentioned, new tools rather than the use of a potential function seem to be necessary. Another challenging direction is to obtain a tight bound on the price of stability, and examine how the price of anarchy can be bounded while fixing other game parameters besides the technological quality. At a higher level, one may consider the partial correlation case, in which a user reacts to a channel state that incorporates both global and local temporal conditions.

REFERENCES

APPENDIX I
PROOFS FOR SECTION III

Proof of Proposition 1. For the proof, we shall make use of the partial derivatives of the aggregate utility, given by
\[
\frac{\partial u(p)}{\partial p^m_i} = \pi_i(R_i^m - \sum_{l \neq m} p_l^i R_l^i(1 - p_l^i)) \quad \text{and} \quad\]
\[
\frac{\partial u(p)}{\partial p^m_i} = \pi_i \rho_i \frac{\lambda_i^m}{\rho_i (1 - p_i^m)} - \sum_{l \neq m} \frac{p_l^i}{1 - p_l^m} R_l^i. \quad \text{(16)}
\]
Let \( \rho_i = \prod_k (1 - p_i^k) \). For any \( p \in \text{int}(E) \), (16) can be rewritten as
\[
\frac{\partial u(p)}{\partial p^m_i} = \pi_i \frac{\rho_i}{1 - p_i^m} \times \]
\[
\left( \frac{R_i^m}{1 - p_i^m} - \frac{\lambda_i^m}{\rho_i (1 - p_i^m)} - \sum_{l \neq m} \frac{p_l^i}{1 - p_l^m} R_l^i \right). \quad \text{(17)}
\]
where we used the fact that \( 1 - p_i^m > 0 \) and \( \rho_i > 0 \) in the interior of the feasible region.

Let \( p \) be an optimal solution of (SWP) and consider some state \( i \). If this state is used by a user with probability \( 1 \) then obviously no other user transmits at this state and the claim immediately follows.

The claim obviously holds if no user utilizes the state. Hence, assume that state \( i \) is partially used by some users. Let \( K_i \subset M \) be the subset of users that partially utilize state \( i \). Let
\[
m \in \text{argmin}_{l \in K_i} \frac{R_l^i}{1 - p_l^i}. \quad \text{(18)}
\]
Since \( p \) is optimal it follows that
\[
\frac{\partial u(p)}{\partial p^m_i} \geq 0, \quad \text{since otherwise the aggregate utility can be improved by decreasing } p_i^m. \]

Substituting (17) in (19) and recalling that \( 1 - p_i^k > 0 \) for every \( k \in K_i \), we obtain that
\[
\frac{R_i^m}{1 - p_i^m} - \frac{\lambda_i^m}{\rho_i (1 - p_i^m)} - \sum_{l \neq m} \frac{p_l^i}{1 - p_l^m} R_l^i \geq 0, \quad \text{(20)}
\]
hence
\[
\frac{R_i^m}{1 - p_i^m} \geq \sum_{l} \frac{p_l^i}{1 - p_l^i} R_l^i \geq \frac{R_i^m}{1 - p_i^m} \sum_{l} p_l^i, \quad \text{(21)}
\]
where the last inequality follows from (18). (21) immediately implies that \( \sum_l p_l^i \leq 1 \).

Proof of Theorem 2. Let \( \hat{p} \) be an optimal solution of (SWP) and define the function \( g^m_p : E^m \rightarrow \mathbb{R} \) as follows:
\[
g^m_p(p^m) \triangleq u(p^m, \hat{p}^m) - u(p^m, \hat{p}^m). \quad \text{(22)}
\]
The function \( g^m_p(-) \) quantifies the change in the aggregate utility, if user \( m \) uses a strategy \( p^m \) instead of \( \hat{p}^m \). Consider the following optimization problem,
\[
\max g^m_p(p^m) \quad \text{s.t. } p^m \in E^m. \quad \text{(23)}
\]
If an optimal solution of this maximization problem is \( p^m \), it follows from the definition of \( g^m_p \) that \( (p^m, \hat{p}^m) \) is an optimal solution of (SWP).

Observe that \( g^m_p(p^m) \) is linear in \( p^m \). Since \( E^m \) is by definition a polyhedron (see (5)), (23) is a linear optimization problem. Therefore, an optimal solution of (23) exists at an extreme point of \( E^m \), and it follows that there exists an optimal solution of (SWP) in which user \( m \) utilizes a threshold strategy.

Note that in the above argument starting from an arbitrary optimal solution of (SWP), we achieve an optimal solution of (SWP), in which all users but \( m \) utilize the same strategy and user \( m \) utilizes a threshold strategy. Repeating the same argument for all users it follows that there exists an optimal solution of (SWP) where all users use threshold policies.

APPENDIX II
PROOFS FOR SECTION IV

Proof for Proposition 2. The idea behind of proof is to construct a NE starting from a threshold multi-strategy \( \hat{p} \). This is done in the 4 steps below. We then compare the aggregate utility of the threshold strategy to that of the obtained NE, by examining the worst case efficiency loss in each step of the construction.

a) Let \( \mathcal{H} \) be the set of states such that each state \( i \in \mathcal{H} \) satisfies \( \hat{p}_i^m > 0 \) for some \( m \in M \). For each \( i \in \mathcal{H} \), define
\[
m_i \in \text{argmax}_{\{k \in M | 0 < p_k^i \leq 1\}} R_k^i - \lambda^k. \quad \text{(24)}
\]

If the set \( \text{argmax}_{\{k \in M | 0 < p_k^i \leq 1\}} R_k^i - \lambda^k \) is not a singleton, \( m_i \) is chosen arbitrarily from the elements of the set. Consider a modified multi-strategy \( q \) of the original multi-strategy \( \hat{p} \), given by
\[
q_k^i = \begin{cases} 
1 & \text{if } i \in \mathcal{H}, k = m_i \\
0 & \text{otherwise}. 
\end{cases} \quad \text{(25)}
\]
Let \( N = \{m_i | i \in \mathcal{H}\} \). Note that from the definition of a threshold policy and (24) it follows that the transmission probability of any user in \( q \) is strictly larger than the transmission probability in \( \hat{p} \) at most for a single state (namely the partially used state, if such exists).

b) The multi-strategy \( q \) can be infeasible, as the power constraint of every \( m \in N \) can be violated by investing extra power in partially used states. Note that if \( k \notin N \), strategy \( q_k^k \leq \hat{p}_k^k \) is feasible. Also \( q_k^i \in \{0, 1\} \) for all \( k \in M, i \in \mathcal{H}, \) and no two users utilize the same state.

Let \( \Delta \hat{P}^m, m \in N \), denote the maximum additional power investment required for user \( m \) to utilize strategy \( q_m^m \) instead of \( \hat{p}_m^m \) (recall that each user partially uses at most a single state). This quantity is obviously bounded by \( \pi_{\max} \), since fully utilizing any state requires at most \( \pi_{\max} \) amount.
of additional power. Set $\Delta \hat{P}^m = 0$ if the strategy of $m$ is already feasible. We next obtain a feasible strategy by modifying $q$.

e) Consider the following optimization problem,

$$
\overline{BR}^k (\hat{p}^k, p^{-k}) = \arg \max_{p^k} u^k (p^k, p^{-k})
$$

subject to

$$
\sum_{i \in H} \pi_i \hat{p}^k_i \leq \sum_{i \in H} \pi_i q^k_i \geq \Delta \hat{P}^k
$$

(26)

$\overline{BR}^k (\hat{p}^k, p^{-k})$ denotes the threshold best response of user $k$ to $p^{-k}$ assuming that the power constraint in the optimization problem is equal to power investment under $\hat{p}^k$. Due to the linearity of $u^k (p^k, p^{-k})$ in $p^k$, the problem becomes a linear optimization problem and a threshold strategy solving (26) always exists. Define

$$
\gamma_k = \max \left\{ j \in H \cup \{ h+1 \} \mid \sum_{i=j}^{h} \pi_i q^k_i \geq \Delta \hat{P}^k \right\}
$$

By convention, assume that $\sum_{i=h+1}^{h} \alpha_i = 0$ for any $\alpha_i$, thus $\gamma_k = h + 1$ for any feasible strategy $q^k$. Consider the following iterative algorithm

1. Set $\bar{M} = \emptyset$, $\bar{p} = q^k$.
2. Choose $k \in \arg \min_{k \in M - \bar{M}} \gamma_k$ (if the set is not a singleton choose an arbitrary $k$ in the set).
3. If $\gamma_k < h + 1$ modify $q^k$ to

$$
\tilde{q}^k_i = \begin{cases} 
q^k_i & \text{if } i < \gamma_k \\
\frac{\sum_{i=\gamma_k}^{h} \pi_i q^k_i - \Delta \hat{P}^k}{\pi_{\gamma_k}} & \text{if } i = \gamma_k \\
0 & \text{if } i > \gamma_k 
\end{cases}
$$

(27)

else set $q^k_i = q_i^k$. Let $w^k = \overline{BR}^k (q^k, p^{-k})$.
4. For any $i \in H$ if $0 < w^k_i < 1$ set $w_i^k = 0$.
5. Set $\bar{p} = (w^k, p^{-k})$, $\bar{M} = \bar{M} \cup \{ k \}$. If $\bar{M} = M$ terminate, else go to step 2.

Let $w$ denote the multi-strategy that is achieved upon termination of the above algorithm. It can be readily seen that $w$ is feasible and all states up to some threshold $i$ are used with probability 1 (each state $i \leq \bar{i}$, $i \in H$ is used by a single user with probability 1), while the remaining states are not used at all. The $0 - 1$ property follows from step 4. The threshold state phenomenon can be easily proved, as otherwise one obtains a contradiction with the optimality of $\overline{BR}^k (\hat{p}^k, p^{-k})$.

d) Let $\bar{p}$ be a NE such that $w$ has the same transmission probability assignment as $\bar{p}$ for states $i \leq \bar{i}$. Such $\bar{p}$ is guaranteed to exist by considering a reduced game where only states $i > \bar{i}$ are considered and the power budgets are given by $\bar{P}^m - \sum_{i=1}^{\bar{i}} \pi_i w^m_i$, $m \in M$.

We next show that the efficiency loss between $\bar{p}$ and any $\bar{p}$ is bounded by some fraction of $u(\bar{p})$, where $\bar{p}$ is the initial optimal threshold solution. To that end, we consider the efficiency loss incurred in the transition from $\bar{p}$ to $\bar{p}$ through the path $\bar{p} \rightarrow q \rightarrow w \rightarrow \bar{p}$.

$\bar{p} \rightarrow q$ : Note that $u(q) \geq u(\bar{p})$, since

$$
u(\bar{p}) = \sum_{i \in H} \pi_i \sum_{l=1}^{\bar{p}} \bar{p}_l^i (R_l^i \sum_{k \neq l} (1 - \bar{p}_l^i) - \lambda_l)
$$

$$
\leq \sum_{i \in H} \pi_i \sum_{l=1}^{\bar{p}} \bar{p}_l^i (R_l^i - \lambda_l)
$$

(28)

$$
\leq \pi_i (R_l^m - \lambda_m) \sum_{l=1}^{\bar{p}} \bar{p}_l^i
$$

$$
\leq \pi_i (R_l^m - \lambda_m) = u(q),
$$

where $\sum_{i \in M} \bar{p}_i^i \leq 1$ and $\bar{p}^m_i \leq \bar{p}_i^m$ for all $m \in M$, $i \in H$. The existence of $\bar{p}_i^m$ for all $m \in M$, $i \in H$ satisfying the first inequality follows by considering the aggregate utility maximization problem for each state $i \in H$ separately and using the fact that at optimal strategy $p$, $\sum_i \bar{p}_i^1 \leq 1$, as Proposition 1 suggests.

$q \rightarrow w$ : For any user $m \in M$ if $p_i^k = 0$ for $k \neq m$ whenever $p_i^m > 0$, then $u^m (p^m, p^{-m}) = \pi_i p_i^m (R_i^m - \lambda_m)$. Hence the payoff is a weighted linear combination of the power invested in different states. Now due to linearity, by assuming: (i) $\sum_{i \in H} \pi_i p_i^m \geq \beta$ and (ii) $\sum_{i=h+1}^{h} \pi_i p_i^m \leq \alpha$ it follows that if user $m$ modifies its strategy $p^m$ to a strategy $\bar{p}^m$ such that transmission probabilities in states $i > j$ for a fixed $j$ are set to zero, then $u^m (p^m, p^{-m}) = u^m (p^m, p^{-m}) \leq \frac{\beta}{\alpha} u^m (p^m, p^{-m})$. This follows since $\alpha$ amount of power which is utilized in lower weights contribute at most $\frac{\beta}{\alpha}$ $u^m (p^m, p^{-m})$ to the user’s payoff.

In step e of the algorithm, observe that modifying actions of a user do not affect the payoff of other users. Let $u^m$ be the initial payoff of user $m$ in this step, then $\sum_{m \in M} u^m = u(q)$. Denote the payoff of user $m$ after step 3 by $\hat{u}^m$, it follows that $\hat{u}^m \geq u^m (1 - \frac{\bar{P}}{P_{min}})$ since for users satisfying $\Delta \hat{P}^m = 0$ the payoff actually increases when playing the best response. Also, for users with $\Delta \hat{P}^m > 0$, at least $P_{min}$ is invested in the system and these users stop investing $\pi_{max}$ amount of power in their worst states in step 3 (as in (27)). Then playing best response, the aggregate utility can only increase and it is larger than $u^m (1 - \frac{\bar{P}}{P_{min}})$.

Similarly, in step 4 of the algorithm, every user $m$ using a state partially invests at least $P_{min}$ amount of power and he stops investing at most $\pi_{max}$ amount of power in his worst states. Denote his final payoff by $\tilde{u}^m$. Then, $\tilde{u}^m \geq \hat{u}^m (1 - \frac{\bar{P}}{P_{min}})^2 \geq u^m (1 - \frac{\bar{P}}{P_{min}})^2$. Since users who do not use any state partially do not modify their policies, it follows that $u(q) \left( 1 - \frac{\bar{P}}{P_{min}} \right)^2 = \sum_{m \in M} \left( 1 - \frac{\bar{P}}{P_{min}} \right)^2 u^m \leq \sum_{m \in M} u^m = u(w)$.

$w \rightarrow \bar{p}$ : Finally, it can be seen that $u(w) \leq u(\bar{p})$. Since $\tilde{p}_i^k = w_i^k$ for $i \leq \bar{i}$, $k \in M$ and the contribution of remaining states to the aggregate utility can not be negative as in this case, at least one user can improve his payoff by setting the transmission probabilities in states $i > \bar{i}$ equal to zero and
this contradicts with the fact that \( \bar{p} \) is a NE. To summarize,
\[
    u(\bar{p})(1 - \frac{\pi_{\text{max}}}{P_{\text{min}}})^2 \leq u(q)(1 - \frac{\pi_{\text{max}}}{P_{\text{min}}})^2 \leq u(\bar{p})
\]
(29)
Hence, \( \frac{u(\bar{p})}{u(q)} \leq \frac{1}{(1 - \frac{1}{\pi_{\text{max}}})^2} \) as the claim suggests.

**Proof for Proposition 3.** We present a parameterized example achieving the PoS lower bound for a given \( Q \). Consider a game with two players \( m \) and \( k \). Let \( Q \) be fixed and define \( j = \lfloor \frac{k}{2} \rfloor + 1 \). Choose \( P_{\text{min}} \) such that \( P_{\text{min}} < \frac{1}{\pi_{\text{max}}} \) and \( \pi_{\text{max}} = P_{\text{min}} Q \). Let \( H = \{1, 2, \ldots, h\} \) and \( h > j \). Consider the system with \( \pi_i = \pi = \frac{P_{\text{min}}}{j} + \epsilon < \pi_{\text{max}} \) for sufficiently small \( \epsilon \) at states \( i \in \{1, 2, \ldots, j\} \) (i.e., the best \( j \) states). Also assume that \( \pi_h = \pi_{\text{max}} \) and the remaining \( \pi_i \) are chosen so that \( \sum_{i \in H} \pi_i = 1 \). Let \( P^m = P^k = P_{\text{min}}, \lambda^m = \lambda^k = \lambda \), where \( \lambda \) will be specified along the sequel. Choose rates as \( (R^m_i, R^m_j, \ldots, R^m_k) = (10 + r_1, 10 + r_2, \ldots, 10 + r_j, \ldots, r_h) \), \( (R^k_i, R^k_j, \ldots, R^k_k) = (\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_j, \bar{r}_j+1, \ldots, \bar{r}_h) \), for \( r_h < \cdots < r_{j+1} < \lambda < r_j < \cdots < r_1 < \delta \) and \( \bar{r}_h < \cdots < \bar{r}_j+1 < \lambda < \bar{r}_j < \cdots < \bar{r}_1 < \delta \) for some \( \delta \).

In this setting, the optimal solution is \( p^k = 0, p^m = (1, \ldots, 1, -\epsilon, 0, 0, \ldots) \) where state \( j \) is the partially used state, and \( \epsilon \) is a function of \( \epsilon \) that satisfies \( \epsilon \rightarrow 0 \) as \( \epsilon \rightarrow 0 \). On the other hand, the best NE satisfies \( p^m_i = 1 \) for \( i < j \) and \( p^k_i = 0 \) for \( i \geq j \) whereas \( p^k_i = 1 \) and \( p^k_i = 0 \) for \( i \neq j \) (where we choose \( \lambda \) such that \( \lambda < \bar{r}_j \)). Now choosing \( \epsilon \) and \( \delta \) sufficiently small (so that the contribution of the terms such as \( r_1 \), and \( \bar{r}_1 \) to the aggregate utility is negligible) the aggregate utility is approximately \( 10\pi\epsilon \) in the central optimum, whereas it is 10\pi(1 - \epsilon) in the best NE, hence
\[
    \text{PoS}(Q) \geq \frac{j}{j - 1} = \left(1 - \frac{1}{\lfloor \frac{k}{2} \rfloor + 1}\right)^{-1}.
\]

**Proof for Proposition 4.** The proof is constructive and follows from an example. Fix \( Q, M \) and consider a symmetric game instance with \( H = \{1, 2, \ldots, h\} \) for \( h > Q^{-1} \). Let \( R^m_i = R_i, \pi_i = \pi_{\text{max}} = \frac{Q}{h}, \lambda^m = \lambda, P^m = \frac{2_{\text{max}}}{Q} \) for every \( m \in M \) and \( i \in H \). Assume that \( \sum_{i \in H} \frac{M-1}{R_i} = h - Q^{-1} \), and \( R_i \gg \lambda \) for every \( i \in H \) (it is always possible to construct such a problem instance for a given \( Q \) by choosing \( h \) and \( \{R_i\}_{i \in H} \) properly. It can be seen that there exists an equilibrium \( p \) for every such game instance which satisfies
\[
    p^m_i = 1 - \frac{M-1}{\lambda R_i} \quad \text{for every } m \in M, \ i \in H,
\]
which yields \( u(p) = \sum_{m \in M} u^m(p) = 0 \) at this equilibrium. Note that the given multi-strategy is feasible since for any \( m \in M \),
\[
    V = \sum_{i \in H} \pi_i p^m_i = 1 - \frac{1}{h} \sum_{i \in H} \frac{M-1}{R_i} = 1 - \frac{h - Q^{-1}}{h} = \frac{\pi_{\text{max}}}{Q} = \bar{p}^m
\]
The aggregate utility at a central optimum is obviously greater than 0, as \( R_i > \lambda \) for every \( i \in H \), leading to an unbounded PoA.

**APPENDIX III PROOFS FOR SECTION V**

**Proof for Lemma 1.** Observe that for any user \( m \in M \), its threshold policies are the extreme points of the feasible region \( E_{m} \). Similarly each threshold multi-policy \( p \) corresponds to an extreme point of the joint feasible region \( E \). The idea behind the proof is to upper bound the number of extreme points of the joint feasible region or equivalently the number of threshold policies in the system (see [5] for the definition of extreme points and related definitions).

In general, a polyhedral region that is a subset of \( \mathbb{R}^n \) and is defined by \( m \) constraints is represented by the polyhedron \( \{x \mid Ax \leq b \} \), where \( A \) is a \( m \times n \) matrix and \( b \in \mathbb{R}^m \) is a constant vector. Now, at any extreme point of this polyhedron, at least \( n \) linearly independent constraints are active, and such constraints uniquely define extreme points, hence there are at most \( \binom{m}{n} \) threshold policies.

In our problem, each user has \( h \) decision variables and a total of \( 2h+1 \) constraints. Hence, in total we have \( M(2h+1) \) constraints and \( Mh \) variables. Thus, the number of threshold policies is bounded by
\[
    \left( \frac{M(2h+1)}{M(h+1)} \right)^{M(h+1)} \leq (2e)^{M(h+1)},
\]
where the first inequality follows from the inequality \( (1 + \frac{1}{n})^n \leq e \).

**Proof for Theorem 4.** By restricting users to threshold strategies, the underlying game becomes a finite game (i.e., the game has a finite action space as Lemma 1 suggests), with a potential function given by (14). As such, the finite improvement property (FIP) in potential games, [10], holds: Any sequence of updates, which results in strict improvement in the utility of the user who is modifying its strategy, is finite. Moreover, each finite improvement path terminates in a NE.

By Lemma 1 the number of threshold strategies is bounded by \( (2e)^{M(h+1)} \). Observing that no multi-strategy can occur more than once during the updates (as the potential strictly increases with each update), this implies that number of updates required for convergence is bounded by \( (2e)^{M(h+1)} \).

By Definition 6, a user who can strictly improve its utility can be found in every \( M \) update periods. Hence, the number of update periods required for convergence is bounded by \( M(2e)^{M(h+1)} \).