Mean field stochastic games for SINR-based Medium Access Control

Hamidou Tembine, Pedro Vilanova, Mohamad Assaad, Merouane Debbah

Mean field stochastic games for SINR-based Medium Access Control

H. Tembine
Ecole Superieure d'Electricite
Supelec
tembine@supelec.fr

P. Vilanova
King Abdullah University of Science and Technology
Saudi Arabia
pedro.guerra@kaust.edu.sa

M. Assaad, M. Debbah
Ecole Superieure d'Electricite
Supelec
{assaad,debbah}@supelec.fr

ABSTRACT

In this paper we study medium access control with interference management. We formulate a SINR-based dynamic channel access as a stochastic game in which the players adapt their retransmission scheme based on their own backoff state. We analyze the asymptotics of the system using mean field dynamics. Both stable and unstable behaviors are illustrated.

Keywords

Mean field, stochastic game, admission control, SINR

1. INTRODUCTION

Many wireless node devices, access points, WiFi-enabled laptops, PDAs, and wireless sensors are deployed throughout offices, streets, campuses, and city environments. All wireless node devices using IEEE 802.11x, 802.15.4 ZigBee, 802.16 WiMAX, and Bluetooth share the IP Multimedia Subsystem (ISM) band in the 2.45 GHz range. Due to the shared spectrum, a wireless group of nodes affects any other other wireless of nodes, and even small wireless node devices can cause strong interference for other node devices. Competitive channel usage causes additional volatility in wireless links, and can lead to critical performance degradation in terms of packet delivery and probability of success. In traditional wireless 802.11x local area networks (LAN), pure Aloha, slotted ALOHA, CSMA (Carrier Sense Multiple Access) and TDMA (Time Division Multiple Access) have been used to mitigate the interference of wireless devices, and manage their medium access control (MAC) so that one user's channel access does not collide with another user's channel access. Because of their simplicity and robustness, these protocols have been widely used to manage both single-range intra-interference and inter-interference which can be a dominant cause of throughput degradation. As the usage of WLANs rapidly increases, nearby access points must compete to access the same channels, generally without any coordination or guiding central authority. Accordingly, network throughput is reduced due to increased packet collisions and high cost due retransmissions and thus energy consumption.

Related work Standard Markov chain models, which have been widely used in IEEE 802.11, very often lead to excessive complications. The authors in [9] have studied power-selection based access control and have shown that if more than three power levels are available to each of the users then correlation mechanisms do not improve the performance of the system.

A mean field approach to MAC protocol have been proposed in [2, 1, 6]. The authors in [1, 10] have pointed out that the validity of the *decoupling assumption* (user's backoff state independence at the asymptotic regime) should be justified, not just by a simple fixed point method but a deep study of the stability of the differential equation is needed. This is because the existence and uniqueness of a rest point does not imply that the dynamics converges to this fixed point and the differential equation may have limit cycles. To construct a cycling behavior of the mean field limit we follow the work in [5] in which the strategies of the players are not considered. We will see in the next sections that the strategies of the players play an important role: *under some strategies, the mean field system is convergent, and under some other strategies the mean field dynamics has a limit cycle (and unstable in the sense of Lyapunov).*

Case of interest In this paper, we consider a mean field stochastic game with finite number of classes (types) and illustrate cycling behavior in an heterogeneous SINR-based MAC protocols in wireless networks. Our contribution can be summarized as follows. To clarify the difference between the existing game formulations with our work, we present both static and dynamic formulation of SINR-based MAC interference management problem and classify them in term of information that are available of the players. We prove mean field convergence of SINR-based access control for specific strategies and channel state distributions and characterize them in term of deterministic differential equations. Considering a specific SINR-based admission control, the mean field limit [10, 8] depends on the control parameters and the uncontrolled mean field limit system can be suboptimal depending on the performance metric. Hence, the control parameters give new insights and help in understanding the behavior of the mean field limit dynamics. We observe that the cycling behavior in the SINR-based MAC protocol can be eliminated by changing the strategy. Finally, a differential population game is formulated at the infinite population limit and mean field equilibria are characterized by backward-forward optimality principle.

The rest of the paper is structured as follows. In the next section we present we present different backoff-state dependent game theoretic formulations. After that we present a mean field convergence result and conduct a detailed analysis for two types. Finally, the proof of the results are given in the Appendix.

2. PRELIMINARIES

The signal to interference plus noise ratio (SINR) channel access model have been widely studied in wireless literature. The SINR takes into consideration the received signal strength, the ambient noise level N_0 and the interference from the users that are active. Let $\mathcal{S}$ be the set of channel states, a subset of an Euclidean space. For a successful reception, the rule requires a certain minimum signal quality threshold β_j i.e each user j has a successful packet delivery if

$$SINR_j(s,a) \geq \beta_j$$

where $s = (h_1, \ldots, h_n) \in \mathcal{S}$ is the channel state profile, a is the action profile and

$$SINR_j(s,a) = \frac{|h_j|^2 \bar{p}_j 1\!\!1_{\{a_j = T\}}}{N_0 + \alpha \sum_{j' \neq j} |h_{j'}|^2 \bar{p}_{j'} 1\!\!1_{\{a_{j'} = T\}}},$$

the term $1\!\!1_{\{a_{j'}=T\}}$ denotes an indicator function for the activities of user j : it is equal to 1 if the link j is active (user j is transmitting "T") and 0 otherwise, $\bar{p_j}$ is the power used by j, α/n is a normalization factor taking in account the system load. Since most of wireless environment are of incomplete information, imperfect measurements and dynamic in nature, we propose different formulation of the SINR-based medium access control problem depending on much information are available to the users.

2.1 Static game formulations

We start by static game formulations where the players are the users, transmitters, nodes etc. There are n players. The set of players is denoted by $\mathcal{N} = \{1, 2, \ldots, n\}$.

Known full channel state: Each player $j \in \mathcal{N}$ knows the vector $s = (h_1, \ldots, h_n)$. Each player can choose its action in the set $\{T, W\}$ where T for "to transmit" and W for "wait". Each player knows the mathematical structure of the payoff functions. The payoff function of player j is given by

$$r_j^1(s,a) = 1\!\!1_{\{SINR_j(s,a) \geq \beta_j\}} - c(a_j)$$

where $a_j \in \mathcal{A} := \{T, W\}$, $c(a_j)$ denotes a power consumption cost, $0 = c(W) < c(T) < 1$. A pure strategy of a player is a function of s that maps to an element of $\mathcal{A}$. We denote this game by $G^1 = (\mathcal{N}, \mathcal{A}, s, (r_j^1(s,.))_{j \in \mathcal{N}})$. A Cournot-Nash equilibrium for a given state s is a configuration such that no player can improve its own payoff by unilateral deviation. It is not difficult to see that this one-shot game G^1 with perfect state monitoring (the full channel state vector is assumed to be known by all the players) has at least one equilibrium.

At this point it is important to mention the constraints imposed by this formulation:

• The exact mathematical structure of the payoff function is assumed to be known by all the players. They are able to compute the requirement element and reasoning to act in the game. These assumptions too demanding in the sense that they may require, for example, lots of signalling, unbounded capabilities and high complexity. In sensor networks, some sensors (robots) may not be able to "compute" complex functions. In some scenarios, one can relax this assumption by considering a dynamic game. In the dynamic game version of G^1, one can exploit the fact that the game G^1 is an aggregative game (no need to know all the actions and states, the aggregative term should be sufficient).

• The assumption of full information is clearly too demanding in many wireless networks scenarios where the topology and network conditions are randomly varying. This assumption may require too much feedback to the players and signalling from receivers to transmitters or between the receivers. The first step in relaxing this assumption is to consider the partial state information.

Partial state information: Here, we assume that each player j knows its own channel state h_j (own-CSI) and the distribution μ_{-j} over the states h_{-j} of the other players. Each player j is able to compute the payoff function given by

$$r_j^2(s_j, \mu_{-j}; a) = \mathbb{E}_{s_{-j} \sim \mu_{-j}} \left[r_j^1(s,a) \mid s_j = h_j, \mu_{-j} \right].$$

Let $\Delta(S_{-j})$ be the set of probability measures over S_{-j}, equipped with the canonical sigma-algebra. Then, $\mu_{-j} \in \Delta(S_{-j})$. Note that r_j^2 is defined over $S_j \times \Delta(S_{-j}) \times \mathcal{A}^n$ where S_j denotes the channel state space of player j. A pure strategy of player j in this game with partial channel state observation is a mapping from the given own-state h_j, and the distribution μ_{-j} to the action set $\mathcal{A}$. We denote the game by $G^2 = (\mathcal{N}, \mathcal{A}, (s_j, \mu_{-j}), (r_j^2)_{j \in \mathcal{N}})$. A pure equilibrium for the game G^2 can be seen as a Bayesian-Nash equilibrium, and it is characterized by $\sigma_j : S_j \times \Delta(S_{-j}) \longrightarrow \mathcal{A}$, and $\sigma_j(s_j, \mu_{-j}) \in \arg\max_{\tilde{a}_j \in \{T,W\}} r_j^2(s_j, \tilde{a}_j, \sigma_{-j})$, $\forall s_j$ and $\forall$j$\in \mathcal{N}$. Note that s_j is own-state and does not necessarily "plays the role" of a type in the classical game theoretic formulations because the channel state is a realization of a certain random variable, not an endogenous element of a player. A second remark is that the consistency relationship between the types and the beliefs should be checked in games with incomplete information.

Structural result for the Bayesian equilibria We remark that the pure best response is a *threshold strategy* in the own channel state: there exists a state h_j^* such that

$$\sigma_j^*(s_j, \mu_{-j}) = \begin{cases} T & \text{if } |s_j| \geq |h_j^*| \\ W & \text{otherwise.} \end{cases}$$

The result is immediate and follows from the fact that the function u_j^2 is monotone in s_j.

None of the state is known: The players do not know any component of the current state. The distributions over the full states are known in addition to the mathematical structure of the payoff functions. Now each player j is able to compute the following payoff function

$$r_j^3(a, \mu) = \mathbb{E}_{s \sim \mu} \left[r_j^1(s,a) \right],$$

where $\mu \in \Delta(\mathcal{S})$. This leads to an expected robust game which we denote by $G^3 = (\mathcal{N}, \mathcal{A}, \mathcal{S}, \mu, (r_j^3)_{j \in \mathcal{N}})$. A pure strategy for player j in the game G^3 corresponds to a function of μ that gives an element of $\mathcal{A}$.

Only the channel state space is known: Now, we assume that the distribution over the state is also unknown but the state space $\mathcal{S}$ is known. Then, each player can adopt different behaviors depending on its way to see the state space. The well-known approaches in that case are the maximin robust and the maxmax approaches (pessimistic or optimistic approaches) and their variants. The payoff of player j in the maximin robust game is given by

$$r_j^4(a) = \inf_{s\in\mathcal{S}} r_j^1(s,a).$$

We denote the game by $G^4 = (\mathcal{N}, \mathcal{A}, \mathcal{S}, (r_j^4)_{j\in\mathcal{N}})$. The equilibria of G^4 are called maximin robust equilibria. Similarly, The payoff of player j in the maximax robust game is given by

$$\bar{r}_j^4(a) = \sup_{s\in\mathcal{S}} r_j^1(s,a).$$

2.2 Dynamic game without backoff state

In this section we formulate dynamic access control games in discrete time. We assume that the channel states are independent of the actions of the players. We distinguish three cases:

1. Stochastic game with perfect state observation: The channel state vector s_t is observed by all the players and instantaneous payoff functions are known. In this case, one can use the classical stochastic game framework to solve the long-run interaction problem.

2. Stochastic game with partial state information: Each player observes a component: Its own channel state i.e $s_j(t)$ is observed by player j. Instantaneous payoff functions are also known. In this case, one can use *partially observable stochastic game* tools and dynamic programming over continuous state space to address the long-term problem.

Each of the above formulations can be with public or private signals, perfect or imperfect monitoring for the actions chosen by the others. Since wireless networks are evolving, the assumption on observability of the other actions may not hold. However, due to the fact that our game has an aggregation structure, a simple signal (message) will be sufficient. Next, we discuss how to deal when no own-CSI are not available at the transmitter side.

3. Stochastic game without state information: The players do not observe the channels (no partial component is observed). The players do not known the state space. They do not know $\mathcal{N}$, they do not know the mathematical structure of the payoff functions. They do not even know if they are in a game. However, they are not in blind environment. We assume that they are able to discover progressively their own action space $\mathcal{A}$ and we assume that each of the players is able to observe an ACK/NACK as it is the case with the carrier sense multiple access (CSMA) with collision avoidance (CA) with ACK/NACK. Now, a private history up to T' is a collection of own-actions and own ACKs up to T'. We denote σ_j^5 a behavioral strategy of player j. We consider the time average payoff of the dynamic game i.e $\liminf_T \mathbb{E}_{s(0),\sigma^5} F^5_{j,T'}$ where $F^5_{T'} = \frac{1}{T'+1}\sum_{t=0}^{T'} r_{j,t}$. In this case, different variants of combined fully distributed payoff and strategy reinforcement learning (CODIPAS-RL) have been developed in [7]. The CODIPAS-RL is a joint and interdependent iterative scheme to learn both payoff function and the associated optimal strategies.

2.3 Stochastic game with backoff state

In the above formulation, the number of retransmission for the same packet was not taken into consideration. In this subsection, we formulate the stochastic game with finite number backoff state per player. Each player j of type θ_j can retransmit at most $K_j < +\infty$ times for the same packet. Assume that each player has a saturated queue. Time is discrete (slotted) and time space is the set of natural numbers. Assume that the backoff slots are synchronized. The backoff process in IEEE 802.11 is governed by a decision process if the duration of per-stage backoff is taken into account:

- every player j in backoff state y_j attempts transmission with probability $\tilde{\sigma}_j^n$ for every time-slot t;

- If $SINR_j(s(t), a(t)) \geq \beta_j$ the player j succeeds and its backoff state $y_{j,t}^n$ changes to 0;

- otherwise, $y_j^n(t)$ changes to $y_j^n(t+1) = y_j^n(t)+1$ modulo (K_j+1). The number of retransmission by player j for the same packet is up to the maximum stage K_j.

Note that, depending on the parameters (channel states, actions, and the thresholds β) several players can be accepted at the same slot if the admission control conditions are met, in that case their state goes to zero and their backoff processes restart with another packet in the queue.

Next, we describe the evolution of backoff process of each player. The backoff state process of each player is influenced by the decision of all the active users (to transmit or not). Depending on the information that are available to the player, the decision process of player j can be based on the number of active users, their backoff states, past decisions and channel states.

In order to have a Markovian backoff state process we ask the decision process to be Markovian (discrete time 1-step Markov process). Thus, the class of Markovian strategies give a Markov decision process with Markov property for states and strategies. Note that these strategies can be time-dependent and congestion dependent (we allow time-inhomogeneous Markov). Each player can count its number of retransmission which corresponds to its own backoff stage. Thus, the strategy of player j at time t depends on n, x_j, t. We denote it by $u^n_{j,x_j}(t)$. Next, we define the transition kernel of the backoff of player j, $Q_j^n(t, s(t), u^n(t)) = [q^n_{j,x_j,x'_j}(t, s(t), u^n(t))]_{(x_j,x'_j)\in\mathcal{X}_j\times\mathcal{X}_j}$ The probability that a player j does not succeed under the strategy u^n is the outage probability i.e

$$q^n_{j,y_j,y_j+1}(t, s(t), u^n(t)) = u^n_{j,y_j}(t)\times$$

$$\sum_{a^n_{-j}} \left(\prod_{j'\neq j} u^n_{j',t}(a^n_{j'})\right) \mathbb{P}\left(SINR_j(s(t), a^n_{-j}) < \beta_j)\right| u^n(t), a^n_j{=}T\right)$$

The probability that the player j moves from backoff state $y_j < K_j$ to 0 is the probability of successful reception i.e

$$q^n_{j,y_j,0}(t, s(t), u^n) = u^n_{j,x_j}(t)\times$$

$$\sum_{a^n_{-j}} \left(\prod_{j' \neq j} u^n_{j',t}(a^n_{j'}) \right) \mathbb{P}\left(SINR_j(s(t), a^n_{-j}) \geq \beta_j)|u^n(t), a^n_j{=}T\right),$$

and $q^n_{j,y_j,y_j} = 1 - q^n_{j,y_j,y_j+1} - q^n_{j,x_j,0}$.

3. MFSG FOR SINR-BASED MAC

In this section, we describe the mean field stochastic game for SINR-based MAC protocol in large-scale wireless networks. Time is discrete ($t \in \mathbb{N}$). There are n players ($n \geq 2$). There is a set of channel states are represented by $s^n(t) \in \mathcal{S} = \mathcal{H}^n$. For every player j, $\mathcal{X}_j$ is its own state space. An individual state has two components as follows: the type of the player and its backoff state. The type is a constant during the game. The state of player j at time t is denoted by $\bar{x}^n_j(t) = (s^n_j(t), x^n_j(t))$ where $x^n_j(t) = (\theta_j, y^n_j(t))$ is the type, and $y^n_j(t)$ is the backoff state of player j at time t. The set of possible types Θ and backoff states is finite. The maximum number of retransmission for a player with type θ is denoted by K_θ. For every player j, $\tilde{A}_j = \{T, W\}$ is the set of actions of that player. $\mathcal{A}_j : \mathcal{S}_j \times \mathcal{X}_j \longrightarrow 2^{\tilde{A}_j}$ is a set-valued map (correspondence) that assigns to each $(s, x_j) \in \mathcal{S} \times \mathcal{X}_j$ the set of actions $\mathcal{A}_j(s, x_j) = \mathcal{A}_j(s_j, x_j) = \mathcal{A}_j(x_j) = \{T, W\}$ that are available to player j. The action of player j at time t is denoted by $a^n_j(t)$. The *global state* of the system at time t is $(s^n_1(t), x^n_1(t), ..., s^n_n(t), x^n_n(t))$. Denote by $a^n(t) = (a^n_1(t), \ldots, a^n_n(t))$ the action profile at time t. The system $(s^n(t), x^n(t))$ is Markovian once the action profile $a^n(t)$ are drawn under Markovian strategies (depends only on t and the current state). We denote the set of Markovian strategies by $\mathcal{U}$. We assume that there is no differentiation in the service and the coefficient $\bar{p}_j, \beta_j$ are identical with the same type. The players are coupled not only via their instantaneous payoff function $r(s^n(t), x^n(t), a^n(t))$ but also via the state evolution $x^n(t)$ i.e the evolution of $x^n_j(t)$ depends on the states and the actions of the other players. We denote by $x^n_j[u](t)$ the state of player j under the strategy u.

Define $M^n[u, m_0](t)$ to be the current population profile under the strategy u starting from m_0 i.e $M^n_x[u, m_0](t) = \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{x^n_j[u](t)=x\}}$. At each time t, $M^n[u, m_0](t)$ is in the finite set $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}^{|\mathcal{X}|}$, and $M^n_x[u, m_0](t)$ is the fraction of players who belong to population of individual state x. For a subset $X_1 \subseteq \mathcal{X}$, define $M^n[u, m_0](t)(X_1) := \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{x^n_j[u](t) \in X_1\}}$. $M^n[u, m_0](t)$ is also called *occupancy measure* under u, m at time t.

Denote by $\mathcal{B}^n_t$ be the set of players that are in the system at time t. Each player such that $j \in \mathcal{B}^n_t$ takes part in a one-shot interaction at time t, as follows. First, each selected player $j \in \mathcal{B}^n_t$ chooses an action $a_{j,t} \in \mathcal{A}(x_j)$ with probability $u(a_j \mid s_j, x_j, t)$ where (s_j, x_j) is the current state of that player. The stochastic array u can be interpreted as the strategy profile of the population. Note that in the standard model of IEEE 802.11 the strategy is fixed and corresponds to the exponentially decreasing retransmission probabilities i.e $u_{\theta,y_\theta} = \frac{1}{2^{y_\theta} b_0}$ for some constant $b_0 > 0$.

Denoting the current set of active players by $\mathcal{B}^n_t = \{j_1, \ldots, j_k\}$, and given the actions $a_{j_1}, ..., a_{j_k}$ drawn by the k players, we draw a new set of individual states $(x'_{j_1}, ..., x'_{j_k})$ and the channel state moves to s' with probability $w^n_{s;s'}$. Under the above assumptions, the transition kernel L^n is invariant by any permutation of the index of the players within the same type if the Markovian strategy satisfy this property. This implies in particular that the players are only distinguishable through their individual state and type. Moreover, this means that the process $M^n[u, m_0](t)$ is also Markovian. Given any Markov strategy and any vector m of $\Delta(\mathcal{X})$, where $\Delta(\mathcal{X})$ is the space of probability distributions over $\mathcal{X}$, the channel state generates an independent process with distribution ω_s. Let $\mathcal{F}^n_t = \sigma(s^n(t'), x^n(t'), a^n(t'),\ t' \leq t)$ be the filtration generated by the sequence of states and actions up to t. The evolution of the system depends on the decision of the interacting players. Given a history $\underline{h}_t = (s^n(0), x^n(0), a^n(0), \ldots, s^n(t) = s, x^n(t), a^n(t))$. $x^n(t+1)$ evolves according to the transition probability $L^n(x'; x, u, s) = \mathbb{P}\left(x^n(t+1) = x' \mid \underline{h}_t\right)$. The marginal of L^n relatively to player j is denoted by $q^n_{x,x'}(u, m, s)$ and the expected probability relatively to the channel state is $\tilde{q}^n_{x,x'}(u, m)$.

3.1 Detailed analysis

In this subsection we analyze in detail the case where the SINR is reduced to $N_0^{-1}\bar{p}_j \mid \bar{h}_{j,t} \mid^2 \mathbb{1}_{\{a_{j,t}=T\}}$ where $h_{j,t} = \bar{h}_j$ corresponds to a good channel state for player j. In this particular case the system can accept only one player at a time slot when the channel state of this particular player is good. Otherwise the player is not accepted due bad channel states or due to interference from the others. Then, the blocking probability is the probability that the corresponding player transmits times the probability that *at least* one of the others transmits at the same slot in the same interference range (same class or other classes). If the current population is observed, a feedback strategy at time t is a function of m i.e a strategy of a generic player j at time t has the form $g(x, n, t, m)$. Denote by $1 - \xi^n_{x_j}(u^n, m, s)$ the probability that *at least* one of the others transmits at the same slot in the same interference range when the generic user backoff is at level $x_j = x$ and the channel state is s. We assume that the probability that the channel state is $\bar{h}$ is $\omega_{\bar{h}}$. Then, the transition kernel of a generic user is detailed in Appendix. We take the probability of retransmission in the form $u^n_x(t) = \frac{u^n_x(t)}{\phi(n)}$. For the channel state $\bar{h}$, one gets

$$\xi^n_x(t, u^n(t), m(t), \bar{h}) = (1 - u^n_x(t))^{nm_x(t)-1} \prod_{x' \neq x} (1 - u^n_{x'}(t))^{nm_{x'}(t)}.$$

We study the process $M^n[u, m_0](t)$ in two different ways represented by the diagram below:

$$\begin{tikzcd}
M^n[u, m_0](t) \arrow[r, "t \longrightarrow +\infty"] \arrow[d, "n \longrightarrow +\infty"] & \varpi^n[u, m_0] \arrow[d, "n \longrightarrow +\infty"] \\
m[u, m_0](t) \arrow[r, "t \longrightarrow +\infty"] & ?
\end{tikzcd}$$

Figure 1: Non-commutativity of the double limit

For a specific scenario of SINR based access control, we show that $M^n[u, m_0](t)$ converges in probability, as n goes to infinity, to $m[u, m_0](t)$. This limit may have cycles when t goes to infinity. Since the limit cycle is different than the invariant measure of the process $\varpi^n[u, m_0]$ as n goes to ∞, we conclude that the two double limits may not coincide. This leads to a non-commutative diagram (figure 1).

The stationary points under u are the zeros of the function $m \longmapsto \tilde{f}(u, m)$. Due to the invariance of the system over the simplex and the regularity of $\tilde{f}$, the system has at least one zero (Poincaré) i.e there exists $m^*(u)$ zero of $\tilde{f}(u, .)$. The question now is to know if such $m^*(u)$ is unique. Since we are considering types, the stationary points can be reduced to a singleton for some value of the strategy. Even in the singleton case, the problem remains because the uniqueness of a stationary point of the ODE does not imply convergence to this stationary point. Consider $\Theta = \{1, 2\}$. In figure 2 we plot the trajectories of m^1, m^{17}, m^0 for the strategy $(\tau = \frac{4}{5})$

$$(u_{1,0}, u_{1,1}, \ldots, u_{1,20}) = (\frac{1}{2400}, \frac{1}{480}, \frac{1}{40}, \frac{\tau}{40}, \ldots, \frac{\tau^{19}}{40}),$$

and

$$(u_{2,0}, u_{2,1}, \ldots, u_{2,20}) = (\frac{1}{3840}, \frac{1}{64}, \frac{1}{64}, \ldots, \frac{1}{64})$$

We denote this strategy by σ_*^1.

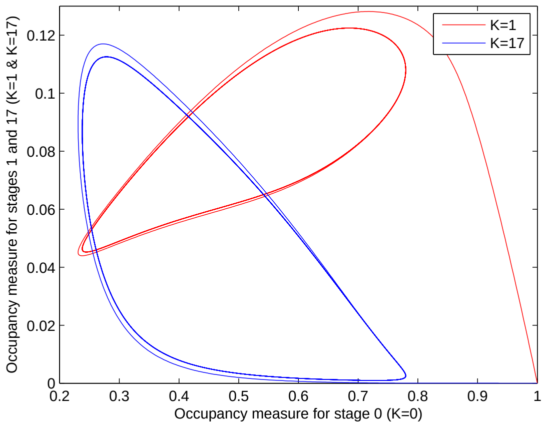

Figure 2: Limit cycle

As mentioned before, the double limit need not to be commutative i.e

$$\lim_n \lim_t M^n[u](t) \neq \lim_t \lim_n M^n[u](t).$$

This phenomenon is in part due to the fact that the stationary distribution of the process ω^n is unique under irreducibility conditions but the dynamics can lead to a limit cycle. As a consequence, many techniques and approaches based on stationary regime (such as fixed-point equation techniques, limiting of frequencies state-actions approaches in sequence of stochastic games, replica methods, interacting-particle systems, statistical independence in large-scale interaction etc) need some justification. The non-commutativity phenomenon suggests to be careful about the use of stationary population state equilibria as the outcome prediction and the analysis of equilibrium payoffs since this equilibrium may not be played. Limit cycles are sometimes more appropriate than the stationary equilibrium approach.

Proposition 1. *The process $M^n_{(\theta,y_\theta)}(nt)$ converges in probability to $m_{\theta,y_\theta}(t)$ which is the solution of the system of ODEs:*

$$\begin{cases} \dot{m}_{\theta,0} = \omega_{\bar{h}} \left[\bar{u}_\theta(t)(1-\gamma(t)) - u_{\theta,0}(t)m_{\theta,0} + u_{\theta,K_\theta}(t)m_{\theta,K_\theta}\gamma(t)\right] \\ \quad +(1-\omega_{\bar{h}}) \left(-u_{\theta,0}(t)m_{\theta,0} + u_{\theta,K_\theta}(t)m_{\theta,K_\theta}\right) \\ \dot{m}_{\theta,y_\theta} = \omega_{\bar{h}} \left[u_{\theta,y_\theta-1}(t)m_{\theta,y_\theta-1}\gamma(t) - u_{\theta,y_\theta}(t)m_{\theta,y_\theta}\right] \\ \quad +(1-\omega_{\bar{h}}) \left(u_{\theta,y_\theta-1}(t)m_{\theta,y_\theta-1} - u_{\theta,y_\theta}(t)m_{\theta,y_\theta}\right), \\ \quad y_\theta \in \{1, ..., K_\theta\}, \theta \in \Theta \end{cases}$$

where $u_{\theta,y_\theta}(t)$ is the strategy of a player from class θ at time t in backoff state y_θ, $\bar{u}_\theta(t) = \sum_{y_\theta=0}^{K_\theta} u_{\theta,y_\theta}(t)m_{\theta,y_\theta}(t)$ and $\gamma(t) = 1 - e^{-\sum_\theta \bar{u}_\theta(t)}$ for $\theta \in \Theta$,.

Here $\bar{u}_\theta(t)$ is the mean field limit of the average attempt rate, $\gamma(t)$ is the blocking probability and $\omega_{\bar{h}}$ is the probability of having a good channel. A detailed proof is given in Appendix.

The equations of the system of ODEs can be intuitively understood. For each channel configuration, the first term and second term on the right-hand side are respectively the inflow caused by blocking probability in the $y_\theta - 1$ backoff stage and the outflow caused by attempts in the y_θ backoff stage times the probability of the corresponding channel state.

The fixed-point equations of the SINR-based access control with several classes of users are obtained by solving the rest point relation of the above system (by letting the right hand side equal to zero, the stationary points give a fixed point relation between γ and v). The *decoupling assumption* or decoupling property consists to say that when the size n goes to infinity, the users becomes mutually independent. Note that this assumption may not hold in general in stationary regime, see the work [5] for IEEE 802.11. This observation concludes the conjectures made in [1, 6] but does not conclude the performance analysis of access control networks. In contrast it opens new questions.

Q1: What are the cases under which the fixed-point equations and the decoupling assumptions are valid?

Clearly, if the ordinary differential equation has a unique global attractor, then the fixed-point equations and the decoupling assumptions are valid, and the diagram (the double limit) is well-defined and is commutative.

The problem is that the rest point (even when it is unique) may not be an attractor. It is because the uniqueness of rest point (also called stationary point, equilibrium states, steady states etc) does not necessarily implies the stability of this stationary point. One may have limit cycle and oscillating behaviors. In this configuration, the rest point is linearly unstable (the Jacobian of the system of ODE is not negative definite at the rest point) for some value of the strategy and the system has a limit cycle that contains the rest point at this relative interior. The phenomenon is related to what is known in evolutionary game dynamics where cycling behavior can be observed. A typical example is the class of Rock-Paper-Scissor games.

The presence of a limit cycle does not allow us to work with the stationary point because the system will never be at this state if the starting point is different. This says that *the statistical independence hypothesis between the users state at the limit* does not hold in stationary regime. Note that the *classical law of large numbers* cannot be used here because even if the users act separately, they are correlated via the interaction. Thus, we need a more elaborated result that exploit the specific structure of our interaction problem to show the convergence. The blocking transmissions from one user is due to a transmission from another user's at the same time slot or it is due bad channel conditions. The blocking transmissions affect the backoff state evolution of both of them. Thus, *the individual states are interdependent.*

Q2: What about the validity of the performance metrics at the limit?

In presence of limit cycle, we suggest to consider the time average performance i.e

$$\lim_{t\to+\infty} \frac{1}{t}\sum_{t'=1}^{t} r_n(t'),$$

where $r_n(t')$ is the instantaneous performance metric at time t' when there are n users. If the limit is not well-defined we replace by $\liminf$ or $\limsup$ in the above definition.

In some scenarios, the time-average of the trajectory of $m(t)$ may converge to the rest point m^* but the convergence does not hold in general. This is because the time-average of the cross-products differs from the product of the time-average of each of them i.e $\lim_t \frac{1}{t}\int_0^t m_x(s)m_{x'}(s)\ ds \neq \left(\lim_t \frac{1}{t}\int_0^t m_x(s)\ ds\right)\left(\lim_t \frac{1}{t}\int_0^t m_{x'}(s)\ ds\right)$. More generally,

$$\lim_t \frac{1}{t}\int_0^t \left(\prod_{l=1}^{k} m_{x_l}(s)\right)\ ds \neq \prod_{l=1}^{k}\left(\lim_t \frac{1}{t}\int_0^t m_{x_l}(s)\ ds\right)$$

The above relation can be seen in some sense as a *violation of propagation of chaos property* at the mean field limit. That is, in order to write the performance along the full trajectory, the convergence needs to be proved. In particular, if one would like to work with stationary distributions of the controlled Markov process or the fixed-point equation then a deep convergence/stability analysis is needed.

Q3: Is it possible to stabilize the system?

The answer to this question relies on the existence of strategy such that the resulting mean field limit is asymptotically stable. We observed that the system can be stabilized by only changing the strategy u to be

$$(u_{1,0}, u_{1,1}, \ldots, u_{1,20}) = (\frac{1}{2400}, \frac{1}{480}, \frac{1}{40}, \frac{\tau}{40}, \ldots, \frac{\tau^{19}}{40}),$$

and

$$(u_{2,0}, u_{2,1}, \ldots, u_{2,20}) = (\frac{1}{3840}, \frac{1}{64}, \frac{1}{64}, \ldots, \frac{1}{64})$$

with $\tau = \frac{1}{2}$ and the linearly stability property of the resulting stationary point can be easily checked by looking the real part of the eigenvalues. We denote this strategy by σ_*^2.

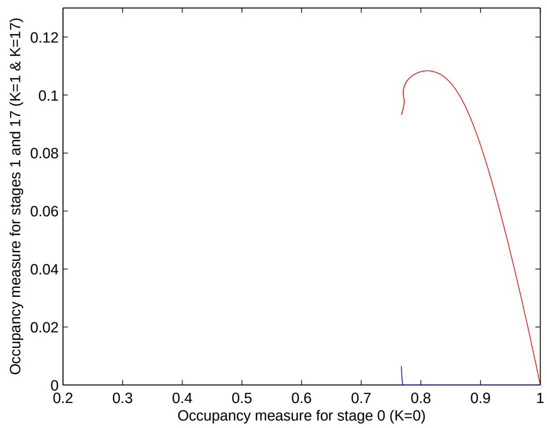

Figure 3: Stable behavior

These observations open several directions. Since we have seen that the system can be globally stable under some strategies, it is natural to ask what is the set of strategies under which the resulting mean field limit has a unique global attractor.

The answer to this question is very important because it leads to the validity domain of the decoupling assumptions and the fixed-point equations. The question $Q3$ is redefined as follows. Let $\mathcal{D}$ be the set of stationary controls u such that the mean field dynamics $\dot{m} = f(u, m)$ is globally convergent.

We are not able to determine completely the domain $\mathcal{D}$ but we have the following results:

• The domain $\mathcal{D}$ is not empty because it contains the size-independent strategy with geometric law. Also the strategy σ_*^2 is in the domain $\mathcal{D}$.

• $\mathcal{D} \subsetneq \mathcal{U}$: $\mathcal{D}$ is a subset of $\mathcal{U}$ but different than $\mathcal{U}$ because σ_1^* is not in $\mathcal{D}$.

Following the same line, we define

$$\mathcal{D}' = \{u \in \mathcal{U},\ \dot{m} = \tilde{f}(u, m)\ \text{ has no limit cycle}\}$$

Since the original problem is to optimize the performance in an autonomous manner, the dynamic optimization problem becomes

$$\sup_{u\in\mathcal{D}'} \text{perf}(u, m).$$

A strategy u^*, solution of this problem give a maximal performance under the non-cycling conditions and stationary strategies. The obtained performance can be less than the value of $\max_{u\in\mathcal{U}} \text{perf}(u, m)$. In order to have a connection between the *stability* and *efficiency*, one can compare the quantity $|\sup_{u\in\mathcal{D}'} \text{perf}(u, m) - \max_{u\in\mathcal{U}} \text{perf}(u, m)|$ which we refer to the *price of stabilization* of the fixed-point analysis.

Remark Our setting can be extended to the path loss case. In place of $|h_j|^2$, we can consider the path loss model: $\frac{1}{d_{j,AP}^{\alpha}}$ where $d_{j,AP} = \| (x_j, y_j, z_j) - (x_{AP}, y_{AP}, z_{AP}) \|_2$ is the distance from the node j located at (x_j, y_j, z_j) to the AP (at (x_{AP}, y_{AP}, z_{AP})), $\alpha \geq 2$ is the pathloss exponent. When the players move around the coverage region, the components x_j, y_j will be random variables.

3.2 Differential population game

In this subsection we provide a mean field equilibrium characterization of the *differential population game* [8] where each generic react to the mean field for a finite horizon $[0, T]$. We first start by a payoff of the form $\bar{r}(u, m)$.

$$(*) \sup_u \ [\bar{g}(m_T) + \int_t^T \bar{r}(u(t'), m(t'))\ dt']$$

subject to the mean field dynamics

$$\dot{m}(t) = m_0 + \int_0^t \tilde{f}(u(t'), m(t'))\ dt'.$$

Definition 1. We say the pair of trajectories $(u^*(t), m^*(t))_{t \geq 0}$ constitutes a consistent mean field response if $u^*(t)$ is an optimal strategy to be above problem (*) where $m^*(t)$ is the mean field at time t and $u^*(t)$ produces the mean field $m[u^*, m_0](t) = m^*(t)$

A consistent mean field response is characterized by a backward-forward equation

$$\left\{ \begin{array}{c} \bar{v}(T, m) = \bar{g}(m) \\ -\partial_t \bar{v}(t, m) = \sup_u \left\{ \bar{r}(u, m(t)) + \langle \nabla_m \bar{v}(t, m), \tilde{f}(u, m) \rangle \right\} \\ m(t) = m_0 + \int_0^t \tilde{f}(u^*(t'), m(t'))\ dt' \end{array} \right.$$

Next, we consider a individual state-dependent payoff $r(x, u, m)$. Define

$$F_T^1(x, u, m) = g(m_T) + \int_t^T r(x(t'), u(t'), m(t'))\ dt'$$

where g is a terminal payoff.

$$(**) \sup_u \ [g(m_T) + \int_t^T r(x(t'), u(t'), m(t'))\ dt']$$

subject to the mean field dynamics

$$\dot{m}(t) = m_0 + \int_0^t \tilde{f}(u(t'), m(t'))\ dt'.$$

Recall that $x(t) = x[u](t)$ is a continuous time jump Markov process under u. We denote by $\bar{q}$ the infinitesimal generator of $x[u](t)$.

Definition 2. We say the pair of trajectories $(u^*(t), m^*(t))_{t \geq 0}$ constitutes a mean field equilibrium if $\{u^*(t)\}_{t \geq 0}$ is a mean field response to be above problem (**) where $m^*(t)$ is the mean field at time t and $u^*(t)$ produces the mean field $m[u^*, m_0](t) = m^*(t)$

PROPOSITION 2. *Consider differential population game problem a single type. Assume that there exists a unique pair (u^*, m^*) such that*

(a) there exists a bounded, continuous differentiable function $\tilde{v}_x : [0, T] \times \mathbb{R}^{|\mathcal{X}|}$, $\tilde{v}_x^*(t, m) = v(t, x, m)$ and differentiable function $m^* : [0, T] \longrightarrow \mathbb{R}^{|\mathcal{X}|}$, $m^*(t) = m[u^*, m_0](t)$ solution to the backward-forward equation:

$$\left\{ \begin{array}{c} v(T, x, m) = g_x(m), \\ -\partial_t v(t, x, m) = \sup_u \left\{ r(x, u, m) + \langle \nabla_m v(t, x, m), \tilde{f}(u, m) \rangle \right. \\ \left. + \sum_{x' \in \mathcal{X}} \bar{q}_{xux'}(m) v(t, x', m) \right\} \\ m(t) = m_0 + \int_0^t \tilde{f}(u^*(t'), m(t'))\ dt' \\ x(t) = x \in \mathcal{X}, m(t) = m, m_0 \in \ \Delta(\mathcal{X}) \end{array} \right.$$

(b) $u^*(x)$ maximizes of the function

$$r(x, u, m(t)) + \langle \nabla_m v(t, x, m), \tilde{f}(u, m) \rangle + \sum_{x' \in \mathcal{X}} \bar{q}_{xux'}(m) v(t, x', m)$$

where $\bar{q}_{xux'}(m)$ is the transition of the infinitesimal generator of $x(t)$ under the strategy u and m, $\sum_{x'} \bar{q}_{xux'}(m) = 0$, the term $\sum_{x' \in \mathcal{X}} \bar{q}_{xux'}(m) v(t, x', m)$ is

$$\sum_{x' \neq x} \bar{q}_{xux'}(m)(v(t, x', m) - v(t, x, m)),$$

$$m[u^*, m_0](t) = m^*(t)$$

Then, $(u^*(t), m^*(t))_{t \geq 0}$ with $m^*(t) = m[u^*, m_0](t)$ constitutes a mean field equilibrium and $\tilde{v}_x^*(t, m^*) = v(t, x, m^*) = F_T(u^*, m^*)$.

Similarly, for multiple types the systems becomes

$$\left\{ \begin{array}{c} v_\theta(T, y_\theta, m) = g_{\theta, y_\theta}(m), \\ -\partial_t v_\theta(t, x, m) = \sup_{u_\theta} \left\{ r_\theta(y_\theta, u_\theta, m) + \langle \nabla_m v_\theta(t, x, m), \tilde{f}(u, m) \rangle \right. \\ \left. + \sum_{y'_\theta} \bar{q}_{y_\theta u_\theta y'_\theta}(m) v_\theta(t, y'_\theta, m) \right\} \\ m^\theta(t) = m_0^\theta + \int_0^t \tilde{f}^\theta(u^*(t'), m(t'))\ dt' \\ y_\theta(t) = y_\theta, m(t) = m, m_0 \in \ \Delta(\mathcal{X}), \theta \in \Theta. \end{array} \right.$$

Note that this result has limited applications because in general the arg max may not be reduced to a singleton. However, for some specific cases such as a particular case of drift limit and strictly monotone payoff, the Proposition can be applied.

Invader strategy Consider now that a new player j enters in the infinite population and adopts a strategy u' different than u. Assume that the (instantaneous) payoff function has the form $r(x(t), u'(t), m[u, m_0](t))$ after taking expectation over the channel states. Note that u' is not in $m[u, m_0]$ because the "effect" of player j is negligible in the infinite population. However, the individual state behavior of player j is influenced by m and u. Namely, the infinitesimal generator is $\bar{q}_{xu'x'}(u, m)$. Then, an optimal response to (u, m) for finite horizon T is a strategy u' such that

$$u' \in \arg\max F_T(u', u, m)$$

where

$$F_T(u', u, m) = \mathbb{E} \left(g(x_j(T), m(T)) + \int_0^T r(x_j(t), u'_j(t), m[u, m_0](t)) dt \mid x_j(0), m_0 \right).$$

In particular a strategy u which is optimal response to itself and which produces m is an appropriate equilibrium configuration. Now, if the fraction of invader increases: the fraction of players which uses u' is size $\epsilon > 0$ and still u is resilient, one gets the notion of RID (resilient to invasion by small fraction of deviants).

4. CONCLUSIONS

In this paper, we have studied mean field asymptotics of SINR-based access control in wireless networks. We have shown that if the strategies are in order of $O(\frac{1}{n})$ and if channel states and SINR-thresholds are well-chosen such that the second moment of the number of backoff state jumps remains bounded then a mean field dynamics is observed. The mean field dynamics may have a limit cycle for some strategy and the cycle disappears for some other strategies. Finally, we have derived mean field equilibrium characterization under specific payoff function at the infinite population limit.

5. REFERENCES

[1] Michel Benaïm and Jean-Yves Le Boudec. A class of mean field interaction models for computer and communication systems. *Perform. Eval.*, 65(11-12):823–838, 2008.

[2] C. Bordenave, D. McDonald, and A. Proutiere. A particle system in interaction with a rapidly varying environment: Mean field limits and applications. *Networks and Heterogeneous Media*, 5:31–62, 2010.

[3] D. L. Burkholder. Martingale transforms. *Ann. Math. Stat.*, 37:1494–1505, 1966.

[4] D. L. Burkholder. Distribution function inequalities for martingales. *Annals of Probability*, 1:19–42, 1973.

[5] J.w. Cho, J.-Y. Le Boudec, and Y. Jiang. On the validity of the fixed point equation and decoupling assumption for analyzing the 802.11 mac protocol. *MAMA*, 2010.

[6] K. Duffy. Mean field markov models of wireless local area networks. *Markov Processes and Related Fields*, 2009.

[7] H. Tembine. Distributed strategic learning for wireless engineers. *Supelec*, 2010.

[8] H. Tembine. Population games in large-scale networks. *LAP ISBN 978-3-8383-6392-9*, 2010.

[9] H. Tembine, E. Altman, R. ElAzouzi, and Y. Hayel. Correlated evolutionarily stable strategies in random medium access control. *International IEEE Conference on Game Theory for Networks, Gamenets*, 2009.

[10] H. Tembine, J. Y. Le Boudec, R. ElAzouzi, and E. Altman. Mean field asymptotic of markov decision evolutionary games and teams. *in the Proc. of GameNets*, May 2009.

APPENDIX

PROOF. of the Proposition 1: To prove the mean field convergence, we use the following steps: To simplify the notations, we will omit the dependency in s at many places.

• Compute the blocking probability for a generic player in each class: Show that this is the probability that the corresponding user's channel is good times the probability that he/she transmits times the probability that *at least* one of the others transmits at the same slot in the same interference range (same class or other classes).

• Denote by $1 - \xi_x^{n,\theta}(u, m, s)$ the probability that *at least* one of the others transmits at the same slot in the same interference range when the generic user backoff is at level s. Then, the transition probabilities $q^{n,\theta}_{x,x'}(u, m, s)$ of the backoff state of a generic user is given by

$$\left\{\begin{array}{c}
\text{for } x = 0 : \\
q^n_{0,1}(t, u^n(t), m(t), \bar{h}) = u^n_0(t)(1 - \xi^n_0(t, u^n(t), m(t), \bar{h})) \\
q^n_{0,0}(t, u^n(t), m(t), \bar{h}) = 1 - u^n_0(t) + u^n_0(t)\xi^n_0(t, u^n(t), m(t), \bar{h}) \\
x \geq 2, q^n_{0,x}(t, u^n(t), m(t), \bar{h}) = 0 \\
\\
\text{for } x \in \{1, \dots, K-1\} : \\
q^n_{x,x+1}(t, u^n(t), m(t), \bar{h}) = u^n_x(t)(1 - \xi^n_x(t, u^n(t), m(t), \bar{h})) \\
q^n_{x,x}(t, u^n(t), m(t), \bar{h}) = 1 - u^n_x(t) \\
q^n_{x,0}(t, u^n(t), m(t), \bar{h}) = u^n_x(t)\xi^n_x(t, u^n(t), m(t), \bar{h}) \\
x' \notin \{0, x, x+1\}, q^n_{x,x'}(t, u^n(t), m(t), \bar{h}, \bar{h}) = 0 \\
\\
\text{for } x = K : \\
q^n_{K,0}(t, u^n(t), m(t), \bar{h}) = u^n_K(t) \\
q^n_{K,K}(t, u^n(t), m(t), \bar{h}) = (1 - u^n_K(t)) \\
x \notin \{0, K\},\ q^n_{K,x}(t, u^n(t), m(t), \bar{h}) = 0 \\
u^n_x(t) = \frac{u^n_x(t)}{\phi(n)} \\
\xi^n_x(t, u^n(t), m(t), \bar{h}) = (1 - u^n_x(t))^{nm_x(t)-1} \prod_{x' \neq x} (1 - u^n_{x'}(t))^{nm_{x'}(t)}
\end{array}\right.$$

• We take the function $\phi(n)$ in order of n. For example, the family of sequences

$$\phi_{\epsilon_1,\epsilon_2,\epsilon_3}(n) = \epsilon_1 n + \epsilon_2 \log(n) + \epsilon_3,\ \epsilon_1 > 0,\ \epsilon_2, \epsilon_3 \geq 0$$

is appropriated. Using the fact that $(1 - \frac{u_{\theta,y_\theta}}{n})^n \longrightarrow e^{-u_{\theta,y_\theta}}$, we can prove that $\xi_x^{n,\theta}(u, m, s)$ has a limit when n goes to infinity. For $u^n = \frac{u}{\phi_{\epsilon_1,\epsilon_2,\epsilon_3}(n)}$ the limit is given by

$$e^{-\sum_x u_x(t) m_x(t)} = e^{-\sum_{\theta \in \Theta} \sum_{y_\theta} u_{\theta; y_\theta}(t) m_{\theta, y_\theta}(t)}.$$

Define block-diagonal the matrix

$$Q^n(u, m, s) = diag[Q^{n,\theta}(u, m, s)]_{\theta \in \Theta};$$

It is easy to see that $n[q^n_{x,x'}(u, m, s) - \mathbb{1}_{\{x=x'\}}]$ has a limit when n goes to infinity. In addition, the resulting function is Lipschitz. We deduce that $nm[Q^n(u, m, s) - I]$ has a limit which is denoted by $f(u, m, s)$. The vector $f(u, m, \bar{h})$ is given by

$$\left\{\begin{array}{c}
f_{\theta,0} = \bar{u}_\theta(t)(1 - \gamma(t)) - u_{\theta,0}(t) m_{\theta,0}(t) + u_{\theta,K_\theta}(t) m_{\theta,K_\theta}(t)\gamma(t) \\
f_{\theta,y_\theta} = u_{\theta,y_\theta-1}(t) m_{\theta,y_\theta-1}(t)\gamma(t) - u_{\theta,y_\theta}(t) m_{\theta,y_\theta}(t), \\
y_\theta \in \{1, ..., K_\theta\}, \theta \in \Theta
\end{array}\right.$$

The vector $\tilde{f}$ is obtained by taking the expectation over the channel state distribution.

• Following similar lines as in the Kurtz's theorem, we show that $M^n(\llcorner nt \lrcorner)$ converges in probability to $m[u, m_0](t)$ and for any $T' > 0$ and $\epsilon > 0$,

$$\mathbb{P}\left(\sup_{t \in [0,T']} \|M^n(nt) - m[u, m_0](t)\| > \epsilon \middle| M^n(0) = m_0, u^n(t) = u \right)$$

goes to 0 when n goes to infinity, where

$$\frac{d}{dt} m_{\theta,y_\theta}(t) = \tilde{f}_{\theta,y_\theta}(u(t), m(t)),\ m(0) = m_0.$$

Since $M^n(.)$ is a discrete Markov decision process over a finite set, we denote by Λ_n the set of all the possible jumps of $M^n(t)$. An element of Λ_n has the form $m'-m$ where $M^n(t)=m$. Then, for a given s, $M^n(t)$ can be written as a function of the jumps:

$$M^n(t) = M^n(0) + \sum_{k=0}^{t-1} \sum_{\lambda_n \in \Lambda_n} \lambda_n 1\!\!1_{\{M^n(k+1)-M^n(k)=\lambda_n\}}.$$

Let $L_{M^n(k),\lambda_n}$ be the probability to have a jump λ_n from $M^n(k)$. Then, $M^n(t)$ can be rewritten as

$$\begin{aligned} M^n(t) &= M^n(0) + \sum_{k=0}^{t-1} \sum_{\lambda_n \in \Lambda_n} \lambda_n 1\!\!1_{\{M^n(k+1)-M^n(k)=\lambda_n\}} \\ &\quad + \sum_{k=0}^{t-1} \sum_{\lambda_n \in \Lambda_n} \lambda_n L_{M^n(k),\lambda_n} - \sum_{k=0}^{t-1} \sum_{\lambda_n \in \Lambda_n} \lambda_n L_{M^n(k),\lambda_n} \\ &= M^n(0) + \sum_{k=0}^{t-1} \varsigma^n(k) + \sum_{k=0}^{t-1} f^n(u, M^n(k)) \end{aligned}$$

where

$$\varsigma^n(k) = \sum_{\lambda_n \in \Lambda_n} \lambda_n [1\!\!1_{\{M^n(k+1)-M^n(k)=\lambda_n\}} - L_{M^n(k),\lambda_n}]$$

and $f^n(u,m,s)$ is the expected drift under the strategy u. Let $\mathcal{F}_t^n = \sigma(M^n(0),\ldots,M^n(t))$. It is clear that the controlled process $z^n(t) = \sum_{k=0}^{t-1} \varsigma^n(k)$ is a martingale with respect to $\mathcal{F}_t^n$.

• Substituting t by $\llcorner nt \lrcorner$ we get

$$M^n(\llcorner nt \lrcorner) = M^n(0) + z^n(\llcorner nt \lrcorner) + \sum_{k=0}^{\llcorner nt \lrcorner -1} f^n(u, M^n(k)).$$

• Using the fact that $nf^n(u,m,s) = nm[Q^n(u,m,s)-I]$ converges to f where the function $m \longmapsto f(u,m,s)$ is locality Lipschitz in m and u, we have that the Cauchy problem associated to the ordinary differential equation (ODE)

$$\left\{ \begin{array}{l} \dot{m}(t) = f(u(t), m(t)) \\ m(0) = m_0 \in \Delta(\mathcal{X}) \end{array} \right.$$

has a unique solution which is $m(t) := m[u, m_0](t)$.

Next, we evaluate the gap between $M^n(\llcorner nt \lrcorner)$ and $m[u,m_0](t)$.

$$\begin{aligned} M^n(\llcorner nt \lrcorner) - m(t) &= M^n(0) - m_0 + z^n(\llcorner nt \lrcorner) \\ &\quad + \sum_{k=0}^{\llcorner nt \lrcorner -1} f^n(u, M^n(k)) - \int_0^t f(u, m(t'))dt' \\ &= M^n(0) - m_0 + z^n(\llcorner nt \lrcorner) + B_{n,t}. \end{aligned}$$

where

$$B_{n,t} = \sum_{k=0}^{\llcorner nt \lrcorner -1} f^n(u, M^n(k)) - \int_0^t f(u, m(t'))dt'.$$

The stochastic term $B_{n,t}$ can be decomposed in three parts:

$$B_{n,t} = \sum_{j=1}^{3} B_{n,t}^j$$

where

$$\begin{aligned} B_{n,t}^1 &= \sum_{k=0}^{\llcorner nt \lrcorner -1} f^n(u, M^n(k)) - \sum_{k=0}^{\llcorner nt \lrcorner -1} \frac{1}{n} f(u, M^n(k)), \\ B_{n,t}^2 &= \sum_{k=0}^{\llcorner nt \lrcorner -1} \frac{1}{n} f(u, M^n(k)) - \sum_{k=0}^{\llcorner nt \lrcorner -1} \frac{1}{n} f(u, m(\frac{k}{n})), \\ B_{n,t}^3 &= \sum_{k=0}^{\llcorner nt \lrcorner -1} \frac{1}{n} f(u, m(\frac{k}{n})) - \int_0^t f(u, m(t'))\, dt'. \end{aligned}$$

We estimate each of the processes $B_{n,t}^j$.

• **Bound for the first term** $B_{n,t}^1$. Let $\bar{\epsilon} > 0$. Using the fact nf^n converges to f, for n sufficiency large, one has

$$\| nf^n(u, M^n(k)) - f(u, M^n(k)) \| \leq \bar{\epsilon}.$$

Thus,

$$\begin{aligned} \| B_{n,t}^1 \| &= \| \sum_{k=0}^{\llcorner nt \lrcorner -1} f^n(u, M^n(k)) - \sum_{k=0}^{\llcorner nt \lrcorner -1} \frac{1}{n} f(u, M^n(k)) \| \\ &\leq \sum_{k=0}^{\llcorner nt \lrcorner -1} \| f^n(u, M^n(k)) - \frac{1}{n} f(u, M^n(k)) \| \\ &= \frac{1}{n} \sum_{k=0}^{\llcorner nt \lrcorner -1} \| nf^n(u, M^n(k)) - f(u, M^n(k)) \| \\ &\leq \frac{1}{n} \bar{\epsilon} \llcorner nt \lrcorner \end{aligned}$$

Hence, $\| B_{n,t}^1 \| \leq \bar{\epsilon} \frac{\llcorner nt \lrcorner}{n},\ \forall t.$

• **Bound for the third term** $B_{n,t}^3$**.** To estimate the term $B_{n,t}^3$, we combine Lipschitz property of the function f relatively to m. Recall that f is c_0-Lipschiz in m if $\forall (m,m')$, $\| f(u,m) - f(u,m') \| \leq c_0 \| m - m' \|$. This implies most linear growth i.e there exists $c_0' > 0$ such that $\| f(u,m) \| \leq c_0'(1 + \| m \|)$, and the Cauchy problem starting at m_0 has a unique solution.

$$\begin{aligned} \| B_{n,t}^3 \| &= \| \sum_{k=0}^{\llcorner nt \lrcorner -1} \frac{1}{n} f(u, m(\frac{k}{n})) - \int_0^t f(u, m(t'))\, dt' \| \\ &\leq \| - \int_{\frac{\llcorner nt \lrcorner}{n}}^t f(u, m(t'))\, dt' \\ &\quad + \sum_{k=0}^{\llcorner nt \lrcorner -1} \frac{1}{n} f(u, m(\frac{k}{n})) - \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(u, m(t'))\, dt' \| \\ &\leq \| \int_{\frac{\llcorner nt \lrcorner -1}{n}}^t f(u, m(t'))\, dt' \| \\ &\quad + \sum_{k=0}^{\llcorner nt \lrcorner -1} \| \frac{1}{n} f(u, m(\frac{k}{n})) - \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(u, m(t'))\, dt' \| \end{aligned}$$

There is $c_3(T) > 0$ such that the first part is less than $\frac{c_3(T)}{n}$ by Lebesgue integrability of f and Lipschitz. The term in the sum of the second part is less than $\frac{c_3'(T)}{n^2}$. By summation, one gets:

$$\| B_{n,t}^3 \| \leq \frac{c_3'(T)}{n^2} \times \llcorner nt \lrcorner + \frac{c_3(T)}{n} \longrightarrow 0.$$

• **Bound for the second term** $B_{n,t}^2$**.** We use the Lipschitz property to estimate the norm of $B^2_{n,t}$.

$$\begin{aligned}\| B^2_{n,t} \| &= \| \sum_{k=0}^{\llcorner nt \lrcorner -1} \frac{1}{n} f(u, M^n(k)) - \sum_{k=0}^{\llcorner nt \lrcorner -1} \frac{1}{n} f(u, m(\frac{k}{n})) \| \\ &\leq \sum_{k=0}^{\llcorner nt \lrcorner -1} \frac{1}{n} \| f(u, M^n(k)) - f(u, m(\frac{k}{n})) \| \\ &\leq \frac{c_0}{n} \sum_{k=0}^{\llcorner nt \lrcorner -1} \| M^n(k) - m(\frac{k}{n}) \|\end{aligned}$$

- Consider the term $\| M^n(\llcorner nt \lrcorner) - m(t) \|$.

$$\begin{aligned}&\| M^n(\llcorner nt \lrcorner) - m(t) \| \\ &= \| M^n(\llcorner nt \lrcorner) - m(\frac{\llcorner nt \lrcorner}{n}) + m(\frac{\llcorner nt \lrcorner}{n}) - m(t) \| \\ &\geq \| M^n(\llcorner nt \lrcorner) - m(\frac{\llcorner nt \lrcorner}{n}) \| - \| m(\frac{\llcorner nt \lrcorner}{n}) - m(t) \|\end{aligned}$$

The expression $\| m(\frac{\llcorner nt \lrcorner}{n}) - m(t) \|$ is in order of $\frac{c_2(T)}{n}$. Suppose that $M^n(0) \longrightarrow m_0$ in probability. Then, for n sufficiently large, $\forall t \in [0, T]$, one has,

$$M^n(\llcorner nt \lrcorner) - m(\llcorner nt \lrcorner) = M^n(0) - m_0 + z^n(\llcorner nt \lrcorner) + B_{n,t}.$$

The norm can be bounded as

$$\begin{aligned}&\| M^n(\llcorner nt \lrcorner) - m(t) \| \leq \bar{\epsilon} + \| z^n(\llcorner nt \lrcorner) \| + \| B_{n,t} \| \\ &\leq 4\bar{\epsilon} + \| z^n(\llcorner nt \lrcorner) \| + \frac{c_0}{n} \sum_{k=0}^{\llcorner nt \lrcorner -1} \| M^n(k) - m(\frac{k}{n}) \|\end{aligned}$$

Next, we estimate the norm of the process $z^n(t)$ by using Burkholder inequality [3, 4] for zero-mean martingales.

$$\begin{aligned}\mathbb{E}\left(\sup_{t \in [0,T]} \| z^n(\llcorner nt \lrcorner) \|^2 \right) &= \mathbb{E}\left(\sup_{[0,T]} \| \sum_{k=0}^{\llcorner nt \lrcorner -1} \varsigma^n(k) \|^2 \right) \\ &\leq c_1(T) \sum_{k=0}^{\llcorner nt \lrcorner -1} \mathbb{E} \| \varsigma^n(k) \|^2\end{aligned} \quad (1)$$

We check that for our SINR-based access control if each player strategy u^n is in order of $O(\frac{1}{n})$ then, we distinguish two classes: (i) number of jumps after a successful transmission, (ii) number of jumps after a collision.

$$\begin{aligned}\mathbb{E} \| \varsigma^n(k) \|^2 &= \mathbb{E} \| \sum_{\lambda_n \in \Lambda_n} \lambda_n [\mathbb{1}_{\{M^n(k+1) - M^n(k) = \lambda_n\}} - L_{M^n(k), \lambda_n}] \| \\ &\leq \| \sum_{\lambda_n \in \Lambda_n} \lambda_n [\mathbb{1}_{\{M^n(k+1) - M^n(k) = \lambda_n\}} \|^2 + \| \sum_{\lambda_n \in \Lambda_n} L_{M^n(k), \lambda_n} \|^2 \\ &\| \sum_{\lambda_n \in \Lambda_n} \lambda_n [\mathbb{1}_{\{M^n(k+1) - M^n(k) = \lambda_n\}} \|^2 + \| f^n(M^n(k)) \|^2\end{aligned}$$

By taking the conditional expectation relatively to $\mathcal{F}^n_t$, each of the terms is in order of $\frac{c_1'(T)}{n^2}$. By combining with (1), we get:

$$\mathbb{E}\left(\sup_{t \in [0,T]} \| z^n(\llcorner nt \lrcorner) \|^2 \right) \leq c_1(T) \frac{c_1'(T)}{n^2} \llcorner nT \lrcorner. \quad (2)$$

This means that on a sample path where the random process $\sup_{t \in [0,T]} \| z^n(\llcorner nt \lrcorner) \| \leq \bar{\epsilon}$ one gets

$$\| M^n(t) - m(\frac{t}{n}) \| \leq 5\bar{\epsilon} + \frac{c_0}{n} \sum_{k=0}^{t-1} \| M^n(k) - m(\frac{k}{n}) \| \quad (3)$$

By taking $\Pi_t = \| M^n(t) - m(\frac{t}{n})) \| \geq 0, \epsilon = 5\bar{\epsilon} > 0, \bar{\mu}_k = \frac{c_0}{n} > 0$ one has, $\Pi_t \leq \epsilon + \sum_{k=0}^{t-1} \bar{\mu}_k \Pi_k$. By (discrete) Gronwall inequality, we get $\Pi_t \leq \epsilon e^{\sum_{k=0}^{t-1} \bar{\mu}_k} = \epsilon e^{\frac{c_0}{n} t}$.

- We deduce that

$$\begin{aligned}&\mathbb{P}\left(\sup_{t \in [0,nT]} \| z^n(\llcorner nt \lrcorner) \| \leq \bar{\epsilon} \right) \\ &\leq \mathbb{P}\left(\sup_{t \in [0,nT]} \| M^n(t) - m(\frac{t}{n}) \| \leq \epsilon + \frac{c_0}{n} \sum_{k=0}^{nT-1} \| M^n(k) - m(\frac{k}{n}) \| \right) \\ &\leq \mathbb{P}\left(\sup_{t \in [0,nT]} \| M^n(t) - m(\frac{t}{n}) \| \leq \epsilon e^{c_0 \frac{\llcorner nT \lrcorner}{n}} \right).\end{aligned}$$

By (2), we know that $\mathbb{P}\left(\sup_{t \in [0,nT]} \| z^n(\llcorner nt \lrcorner) \|^2 \leq \bar{\epsilon} \right) \longrightarrow 1$ when $n \longrightarrow +\infty$. Hence,

$$\mathbb{P}\left(\sup_{t \in [0,nT]} \| M^n(t) - m(\frac{t}{n}) \| > \epsilon e^{c_0 \frac{\llcorner nT \lrcorner}{n}} \right) \longrightarrow 0.$$

This completes the proof. □

PROOF. sketch Proof of Proposition 2: If $m^*(t) = m^*(t; u^*(t))$ solution of the mean-field limit dynamics which is substituted into the Hamilton-Jacobi-Bellman (HJB) equation, the differential population game results in the solutions of a novel HJB equation given by

$$\tilde{f}^* = \tilde{f}(u, m^*(t)), r^* = r(., u, m^*(t)).$$

Since the new PDE admits a solution and the control $u^*_x(t) = u_x(.)$ optimizes the righthand side is a best response to m^*; this means that the optimal response of the individual player generates a mean-field limit which is a solution of ODE (thus, consistent) and the players compute their controls as a function of this mean-field, it follows that $(u^*; m^*)$ is a mean-field equilibrium □