Age-Dependent Distributed MAC for Ultra-Dense Wireless Networks

Dheeraj Narasimha
School of Electrical,
Computer and Energy Engineering
Arizona State University, Tempe
ddnarasi@asu.edu

Srinivas Shakkottai
Department of Electrical
and Computer Engineering
Texas A&M University, College Station
sshakkot@tamu.edu

Lei Ying
Electrical Engineering
and Computer Science
University of Michigan, Ann Arbor
leiying@umich.edu

IEEE INFOCOM 2021 - IEEE Conference on Computer Communications | 978-1-6654-0325-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/INFOCOM42981.2021.9488674

***Abstract*—We consider an ultra-dense wireless network with N channels and $M = N$ devices. Messages with fresh information are generated at each device according to a random process and need to be transmitted to an access point. The value of a message decreases as it ages, so each device searches for an idle channel to transmit the message as soon as it can. However, each channel probing is associated with a fixed cost (energy), so a device needs to adapt its probing rate based on the "age" of the message. At each device, the design of the optimal probing strategy can be formulated as an infinite horizon Markov Decision Process (MDP) where the devices compete with each other to find idle channels. While it is natural to view the system as a Bayesian game, it is often intractable to analyze such a system. Thus, we use the Mean Field Game (MFG) approach to analyze the system in a large-system regime, where the number of devices is very large, to understand the structure of the problem and to find efficient probing strategies. We present an analysis based on the MFG perspective. We begin by characterizing the space of valid policies and use this to show the existence of a Mean Field Nash Equilibrium (MFNE) in a constrained set for any general increasing cost functions with diminishing rewards. Further we provide an algorithm for computing the equilibrium for any given device, and the corresponding age-dependent channel probing policy.**

I. INTRODUCTION

In recent years, the number of smart wireless devices has exploded. Smart wireless devices are used in smart homes, self-driving cars, and are an integral part of the Internet of Things (IoTs). As these devices proliferate, so does the demand on the spectrum. As a result, it becomes an increasingly important problem to allocate channels (frequency bands) to these devices in an efficient way. For many applications, centralized solutions are increasingly infeasible because of the dynamical nature of the systems (devices may continuously join/leave the system), resulting in a need for distributed algorithms that maintain efficiency even as the network scales.

We consider the following scenario. Status messages with fresh information are generated by the devices and need to be communicated to a central node or peers or a base station. This is typically seen in IoT systems including surveillance using wireless sensor networks and V2V/V2X communications in autonomous driving. In these systems, old messages become redundant when a new message is generated and thus may be dropped. It is also critical to send the messages as soon as possible. Status messaging plays a vital role in many other networks including wireless sensor networks during aggregation of sensor information. Status messaging is likely to play an important role in control channels between central access points and devices especially in 5G networks. However, managing medium access in a system with a very large number of devices, e.g., providing situational awareness across many ongoing processes in a factory, posses a significant challenge.

We consider the question of distributed medium access control (MAC) in this regime of a very large number of sensors providing situational updates at a base station. Each sensor generates samples at random times, and only the most recent sample at the sensor is a candidate for transmission (older ones are dropped). As this sample ages, its value goes down correspondingly. The sensor must probe to find an idle channel to transmit its message, and such probing costs it energy. Thus, each sensor must trade-off age of the current sample and probing rate in order to ensure highest value of its situational update. But since the sensor needs to estimate the probing rates of the other sensors to determine this value, the complexity of decision making increases exponentially with the number of sensors in the system. How are we to analyze such a system?

An attractive means of handling the complexity of this setting is to consider the mean field regime, wherein the assumption is that each sensor node uses a state-dependent sensing policy that only depends on the other sensors' states via a belief about the steady state fraction of busy channels engendered by their actions. Validity of this model depends on the existence of such a steady state distribution—the mean field limit—for a given policy employed by all sensors as the number of sensor nodes becomes asymptotically large.

Assuming that the mean field limit exists for a given class of policies, one can then ask about the nature of an equilibrium policy in the game setting, wherein each sensor node is considered as a strategic agent that attempts to maximize its own payoff. The strategic model is natural in the case of different sensors measuring diverse parameters (eg., temperature, location, inventory, etc.), with each sensor wanting to ensure high-fidelity situational awareness regarding the parameter that it is responsible for sensing. The equilibrium concept here is a Mean Field Nash Equilibrium (MFNE), under which the policy applied by all agents generates a mean field distribution, and none of the agents has an incentive to deviate from this policy given its belief about the mean field distribution.

Main Results

Under our problem formulation, each device competes with other devices over a common set of channels that they must share. A device needs to solve an infinite horizon MDP to maximize its discounted sum of rewards, given its age-dependent state and the fraction of busy channels γ. We formulate the problem as a mean field game wherein the device uses the mean field limit to approximate the fraction of busy channels. A summary of our main results is as follows:

- **Existence of Mean Field Limit (MFL).** We begin by showing that the mean field limit exits, i.e., there is a class of policies under which the steady state distribution of the nodes converges to a fixed distribution as the number of nodes becomes large. In Theorem 1, we show that the system converges to the mean field limit under the homogeneous setting where all devices employ the policy $\alpha := \{\alpha_0, \alpha_1....\}$, where α_i is the probing rate when the age of the message is i, as long as the policy belongs to a specific class of policies $\mathcal{P}$. Therefore, we may now assume that each device attempts to solve the MDP problem to obtain the policy α when the fraction of busy channels γ in the mean field limit is given.
- **Age-based probing policy.** We next characterize the optimal policy in the game setting as a function of the countably infinite age set. We show that when reward of transmitting a message with age i, denoted by R_i converges to zero as $i \to \infty$, we can characterize the value functions $\{u_i\}$ as a decreasing sequence, i.e., value in non-increasing in age. We conclude that the optimal policy $\{\alpha_i\}$ exists and converges in the limit as age i goes to infinity, i.e., it is well defined for all values of age. We use these two facts to show that we can approximate the infinite dimensional vector $\alpha = \{\alpha_0, \alpha_1, ...\}$ with some policy in $\mathcal{P}$. These results show that the mean field model is consistent with the set of policies we choose.
- **Existence of MFNE.** Having shown that the mean field limit exists for some class of policies and by characterizing the set of policies for the system at the mean field limit, we further prove that a fixed point exists using Brower's fixed point theorem. These fixed points are Mean Field Nash Equilibria (MFNE), with devices being restricted to using the space of policies $\mathcal{P}$. We prove this by showing that the map from the space of policies to the fraction of busy channels and the map from the fraction of busy channels to the space of policies are both continuous in Theorem 2.
- **Performance Evaluation of the Age-Dependent Distributed MAC.** In the last section, Section VII, we describe a tractable method to find the MFNE for any given device. We provide simulation results comparing our protocol to the MFG-based protocol D-MAC presented in [1] that optimizes for a throughput-dependent reward for a varying arrival parameter λ. We compare the delay experienced by a packet transmitted at steady state under both protocols under an appropriate choice of parameters to ensure that the policy is well defined for both protocols. Our algorithm, entitled Age-Dependent MAC (AD-MAC) experiences better delays for all values of λ while providing comparable congestion of channels.

II. Related Work

There have been recent advances in the modeling and incorporation of *strictly deadline-constrained stream of requests* with long-term *drop-rate* requirements [2]–[7]. There is also a stream of work emerging on the *Age of Information (AoI)*, that seeks to ensure that the samples received at an aggregation point satisfy constraints on the difference between the current time and generation time of the last received sample, known as "age" [8]–[11]. In each of these cases, the authors assume an exponential rate of arrivals, and service is assumed to be exponential. Our assumptions on the arrivals and transmission rate reflect these assumptions. The key differences between this literature are that in our work, the sensor node is concerned with only the most recent sample (we measure age and the delay cost at sender), and we consider the mean field regime with a large number of sensors.

The mean field approach to study an M-particle, N-dimensional continuous time Markov chain has been investigated extensively in recent years. The key idea is to use an Ordinary-Differential-Equation (ODE) to model the system which can approximates the system in the limit of $M \to \infty$. In such cases, the fixed point of the ODE can be used to approximate the steady state of the Markov chain and the fixed point is called the mean field equilibrium (MFE).

For communication networks, the mean field approach was first used to model MAC protocols in the seminal work [12] to establish the steady-state performance for 802.11 MAC. Later, the performance of 802.11 MAC in the unsaturated case was characterized using the mean field approach in [13]. However, these existing works consider the case of a single interference channel or an interference graph, instead of multi-channel systems considered in this paper.

The mean field approach has been exploited in the game setting famously in [14] which studied a one-shot game under which the asymptotic independence property was used to greatly simplify the decision making. More recent papers have considered repeated games under a variety of different applications [15]–[17]. These papers typically use almost sure convergence of the system to a mean field to simplify the analysis and identify the mean-filed Nash equilibrium based on mean field. Propagation of chaos ([18]–[21]) plays a crucial role in these cases.

This paper relies on Stein's method [22]–[25] instead of propagation of chaos to prove the convergence to mean field limits. Stein's method was crucial in [1] to prove the existence and convergence of the MFNE because it can not only be used to prove convergence, but also can provide the rate of convergence, which becomes necessary to show that the MFNE is an ϵ-Nash-equilibrium. The problem studied in [1] focuses on throughput and does not consider the age of the message. In other words, the reward of transmitting a message is a constant and is independent of the freshness of the message. We greatly generalize the results from [1], providing a framework to solve a subclass of countable state, density dependent Mean field games. Technically, instead of solving a single optimization problem at each stage as in [1], each device must solve a countably infinite state, infinite horizon MDP problem under an arbitrary continuous increasing cost function.

III. System Model and Mean Field Game

System Parameters We consider an N-channel, M-device ultra dense wireless network. We consider the case when both M and N tend to infinity and M/N is a constant. Each device in the network generates status messages following a Poisson process with rate λ. The age of a message at each device evolves following an exponential clock that ticks with rate $1/\delta$. We call this clock the delay clock. When the delay clock ticks, the age/delay of the message increases from i to $i+1$. **States** The states in our Markov chain are indexed from $\{-2, -1, 0, 1, ...\}$. When the device has no message to transmit, we define its state to be -1. The state -2 indicates that the device is currently transmitting the message. When a new message arrives, the device goes to state 0 indicating that the message has no delay (zero age). The state $i > 0$ indicates non-zero age. Therefore, the state of the device changes from i to state $i+1$ with rate $1/\delta$ following the delay clock.

Control Policy Each device maintains a separate exponential clock with rate $\alpha_i \in (0, A)$ for some constant A when in state i. When this clock ticks, the device will probe one channel at random to see if it is free. If the channel is free, the device grabs the channel and starts transmitting its message. In this case, the state of the device moves from i to state -2 which indicates that it is in the transmitting state. If a new message arrives when the device is in the probing state, the age of message is reset to 0, which means that the state of the device transits from state i to state 0. On the other hand if a new message arrives while the device is transmitting its message, the message is stored and transmitted immediately after the current message without giving up the channel. The state space diagram for this system in Fig 1.

Reward and Cost function A device receives a reward of R_i for transmitting a message with age i (i.e. at state i). It is important for devices to transmit their messages with fresh information, so the reward R_i decreases in i and decreases to 0 as i increases. Further, we assume that for messages that arrive while the device is in state -2, the reward for transmitting each of these messages is a constant r_{-2}. Additionally, with probing rate α, the device needs to pay a cost of $\hat{c}(\alpha)$, which is a strictly increasing function with $\hat{c}(0) = 0$. Consider the corresponding jump process for the CTMC described above, then we use t_j to denote the time between the j^{th} tick and the $j+1^{st}$ tick of the overall exponential clock. Note that with probing rate α, the expected transition time is $\frac{1}{\alpha+\lambda+\frac{1}{\delta}}$.

Bellman Equation Given the above model, each device maximizes the following discounted infinite-horizon problem:

$$u_x = \frac{1}{A+\lambda+1/\delta} \times \max_{\alpha\in[0,A]} \left(\alpha+\lambda+\frac{1}{\delta}\right) E\left[\sum_{j=0}^{\infty} \beta^j \left(\mathbb{1}_{X(j+1)=-2} R_{X(j)} - \mathbb{1}_{\{X(j+1)=X(j) \bigcup X(j+1)=-2\}} \hat{c}(\alpha_{X(j)}) \right) \middle| X(0)=x \right], \tag{1}$$

where β is the discount factor, $X(0)$ is the initial state of the CTMC, and $X(j)$ is the state of the CTMC after the j^{th} transition. We can view this Bellman equation as the normalized time averaged reward that each device obtains when initialized with state x with a constant $(A+\lambda+1/\delta)$. If one imagines each device to maintain a super clock used to simulate all the events, then this super clock will need to have a tick rate of $(A+\lambda+1/\delta)$. One can therefore view $\frac{1}{A+\lambda+1/\delta}$ as a normalized unit of time. The time spent in state $X(j)$ before the next event is given by $\frac{1}{\alpha+\lambda+1/\delta}$ with probing clock ticks with rate α.

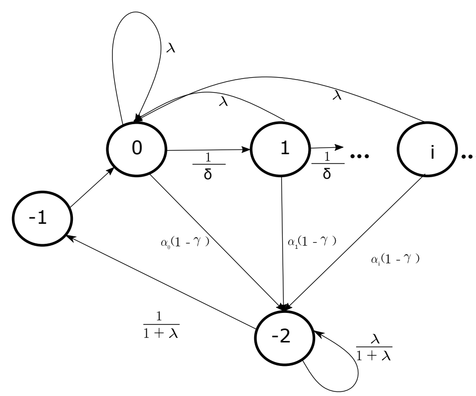

Figure 1. State space model for the Markov chain

Note that, in order for a given device to find the optimal policy it must take into account the fraction of busy channels, $\gamma(t)$. Which itself is a density dependent random process that is determined by the states of all the devices. In other words, this fraction of busy channels $\gamma(t)$ couples all M devices, and makes it intractable to solve the steady-state of the system and the optimal policy based on the Bellman equation.

A. Mean Field Game

To overcome this difficulty, we approach the problem from a mean field game perspective. We assume the time-scale separation (a similar assumption used [1]) such that the devices adapt their policies in a slower time-scale than the convergence of the system to its steady-state with a fixed policy.

Under this time-scale separation, with a fixed policy α for all devices, the steady-state of the stochastic system converges to the equilibrium point of a mean field model, called the mean field-limit, to be defined in Section IV as the system size increases. In particular, the fraction of busy channels $\gamma(t)$ converges to a point mass γ, i.e. a constant. Let us denote this mapping under this mean field limit when the policy α is given by T_1 such that

$$T_1 : \alpha \to \gamma.$$

This occurs at a *fast* time-scale.

Now, for the fixed γ under the previous policy, each device solves the Bellman equation to determine a new policy for the given γ. Thus, for a fixed γ, each device finds its policy based on the Bellman equation with a constant γ (see details on the structure of the policy in Section V). We denote this mapping by T_2 such that

$$T_2 : \gamma \to \alpha.$$

We now define an Mean Field-Nash-Equilibrium (MFNE) as a policy, α^*, such that

$$\alpha^* = T_2(T_1(\alpha^*)),$$

i.e., a fixed point of mapping $T_2(T_1(\cdot))$. In order to show the existence of such a fixed point, we will need to show that the composition of maps T_1 and T_2 is a continuous function. We have already characterized T_1, and in the following sections, we will characterize T_2 to show that both T_1 and T_2 are continuous. Hence, the composition of the two maps is also continuous. The existence of a fixed point follows from Brouwer's fixed point theorem. Brouwer's fixed point theorem requires the map from some set $\Omega \to \Omega$ to be continuous and the set Ω must be closed and compact. The last condition can be checked easily, we will prove continuity in Section VI after proving the convergence to the mean field limit in Section IV and analyzing the policy structure in Section V.

IV. Mean Field Limit Under a Given Policy: The mapping T_1

This section focuses on the convergence of the stochastic system to its mean field limit when the policy $\alpha := \{\alpha_0, \alpha_1 ...\}$ is chosen by each device is fixed and such that there exists a finite K such that $\alpha_k = 0$ for all $k > K$. Let $\mathcal{P}^K$ denote the set of all such policies. We will present the necessary assumptions in Section V so that this condition will be satisfied. Let $Q_j(\infty)$ denote the number of devices in state j at the steady state, so $Q_j(\infty)/M$ is the fraction of devices in state j. We further denote S_K to be the fraction of devices who are in state j such that $j \geq K$, i.e., the fraction of devices with delay greater than or equal to K.

In the limit as N and M go to infinity, we will show that the fraction of devices in state $Q_j(\infty)/M$ $j = 1, \cdots, K-1$ and S_K converge weakly to π_j where π_j is the equilibrium point of the mean field model below:

$$\begin{aligned}
\frac{dq_{-1}}{dt} &= -\lambda q_{-1} + \frac{1}{1+\lambda} q_{-2} \\
\frac{dq_{-2}}{dt} &= \sum_{i=0}^{\infty} (1-\gamma)\alpha_i q_i - \frac{1}{1+\lambda} q_{-2} \\
\frac{dq_0}{dt} &= \lambda\big(1 - q_0 - q_{-2}\big) - q_0 \frac{1}{\delta} - \alpha_0(1-\gamma) q_0 \\
\frac{dq_j}{dt} &= \big(q_{j-1} - q_j\big)\frac{1}{\delta} - \alpha_j (1-\gamma) q_j - \lambda q_j \\
& \qquad j = 1, \cdots, K-1 \\
\frac{ds_K}{dt} &= q_{K-1}(t)\frac{1}{\delta} - \lambda s_K(t)
\end{aligned} \tag{2}$$

The proof is an application of Theorem 1 in [22]. The theorem states five conditions that are sufficient to guarantee the weak convergence to the fixed point of the mean field model. We next verify these conditions under our model:

- **Bounded transition rate:** This condition can easily be verified from the system model. At any point in time, the rate of transition from any state to any other state is bounded above by $A + \lambda + 1/\delta$.
- **Bounded state transition condition:** Since our model is a collection of M CTMCs whose transition rates are determined by exponential clocks, at most one transition can occur at a time. Therefore, the state transitions are bounded.
- **Perfect Mean Field Model:** Using the system model it can be checked that the equations (2) are derived from the detailed balance equations.
- **Partial Derivative condition:** It can be checked that the partial derivatives for the system (2) exist and are Lipschitz.
- **Stability conditions:** Let $\{\pi_i\}$ be the equilibrium point for the dynamical system given by (2). Then we use $\{\epsilon_i\}$ to denote the difference, $\pi_i - q_i$. One can check that the system is globally exponentially stable by using the following Lyapunov function:

$$V := \sum_{i=0}^{K-1} |\epsilon_i| + |\epsilon_K| + |\epsilon_{-1}| + |\epsilon_{-2}|$$

This proof is highly technical but space consuming and does not add much to the discussion of the paper itself, we thus have chosen to omit it due to space constraints.

The following theorem summarizes the result that the Markovian system, $\left(\frac{Q_1(\infty)}{M}, \cdots, \frac{Q_{K-1}(\infty)}{M}, S_K(\infty)\right)$, converges weakly to the fixed point of the dynamical system above,

Theorem 1. *If every device follows a fixed policy α defined in the beginning of this section, then the stationary distribution of the system converges to the unique equilibrium point of system (2). This defines the mapping $T_1 : \alpha \to \gamma$.*

V. Characterizing the policy

If γ denotes the fraction of busy channels, which (under the mean field model) remains to be a constant and is known to a device, then the Bellman equation (1) for the discounted problem becomes

$$u_i = \max_{\alpha\in[0,A]} \frac{\alpha+\lambda+1/\delta}{A+\lambda+1/\delta}\Bigg(\frac{(1-\gamma)\alpha}{1/\delta+\lambda+\alpha}\big(R_i+\beta u_{-2}\big) - \hat{c}(\alpha)\frac{\alpha}{\alpha+\lambda+1/\delta} + \beta\frac{\gamma\alpha}{1/\delta+\lambda+\alpha}u_i + \beta\big(\frac{1/\delta}{1/\delta+\lambda+\alpha}u_{i+1}\big) + \beta\big(\frac{\lambda}{1/\delta+\lambda+\alpha}u_0\big)\Bigg). \quad (3)$$

We will henceforth refer to $\hat{c}(\alpha)\frac{\alpha}{\alpha+\lambda+1/\delta}$ as $c(\alpha)$, which obeys all the properties of $\hat{c}(\alpha)$. Note that conditioned on a state transition occurs when the device is in state i, we have the following possibilities:

- With probability $\frac{(1-\gamma)\alpha}{1/\delta+\lambda+\alpha}$, the probing clock ticks and the device finds an idle channel. In this case, the device pays a cost $c(\alpha)$ and receives a reward R_i. The device transits to state 2.
- With probability $\frac{\gamma\alpha}{1/\delta+\lambda+\alpha}$, the probing clock ticks and the device fails to find an idle channel. In this case, the device pays a cost $c(\alpha)$. The device remains in state i.
- With probability $\frac{\delta}{1/\delta+\lambda+\alpha}$, the age of the message increases by one. In this case, the device moves to state $i+1$.
- With probability $\frac{\lambda}{1/\delta+\lambda+\alpha}$, a new message arrives and replace the current message in waiting. In this case, the device moves to state 0.

Note that the term u_i appears at both sides of the Bellman equation, by combining the two terms, we have (this is relatively easy to verify and is omitted due to space constraints):

$$u_i = \max_{\alpha\in[0,A]} \frac{1}{1/\delta+\lambda+A-\gamma\beta\alpha}\Big\{(1-\gamma)\alpha\big(R_i+\beta u_{-2}\big) - c(\alpha)\left(\alpha+\lambda+1/\delta\right) + \frac{\beta}{\delta}u_{i+1} + \beta\lambda u_0\Big\}$$

with the special cases

$$u_{-1} = \beta\frac{\lambda}{A+\lambda+1/\delta}u_0$$

and

$$u_{-2} = \frac{1+\lambda}{A+\lambda+1/\delta}\Bigg(r_{-2} + \beta\frac{1}{1+\lambda}u_{-1} + \beta\frac{\lambda}{1+\lambda}u_{-2}\Bigg).$$

Subtracting both sides of the previous equation by $\beta\frac{\lambda}{1+\lambda+A}u_{-2}$, multiplying throughout by $\frac{1+\lambda}{1+\lambda(1-\beta)}$ and substituting u_{-1} in terms of u_0, we obtain

$$u_{-2} = r_{-2}\frac{1+\lambda}{1/\delta+\lambda(1-\beta)+A} + \beta^2 u_0\frac{\lambda}{1/\delta+\lambda(1-\beta)+A}.$$

Now define

$$r_0 := r_{-2}\frac{1+\lambda}{1/\delta+\lambda(1-\beta)+A}$$

and

$$\eta := \beta\frac{\lambda}{1/\delta+\lambda(1-\beta)+A},$$

which gives us the following expression for u_{-2},

$$u_{-2} = r_0 + \eta\beta u_0$$

Note that we have essentially treated the fraction of busy channels as a constant in studying the Bellman equation above. In other words, a device optimizes its probing strategy assuming γ is fixed. This assumption is justified so long as the system converges to a point mass given by the mean field limit, i.e, satisfies the conditions of Theorem (1).

Before we proceed, we make the following remarks that will be helpful in later sections.

Remark 1.

1) u_i is bounded below by 0. The lower bound is achieved when we choose $\alpha = \{0, 0, ...\}$, i.e. to do nothing at all no matter the delay, reward or cost.
2) u_i is bounded above by $\frac{R}{1-\beta}$ with $R = R_0 + r_{-2}$.

Now, assume we begin by initializing all the devices with the same policy in $\mathcal{P}^K$. From the previous section, it is clear that the fraction of busy channels will converge weakly to some fixed γ. In the rest of the section, we will show that the sequence of value functions $\{u_i\}$ is decreasing and so is well defined for all $i \in \{-2, -1, 0, 1, ...\}$. Since both $\{u_i\}$ and $\{\alpha_i\}$ are infinite sequences, we need to establish that $\{\alpha_i\}$ is well defined in the limit as i goes to infinity (it is not immediate that the map from $\{u_i\}$ to $\{\alpha_i\}$ is sequentially continuous). We use the convergence of $\{u_i\}$ to show that the sequence of $\{\alpha_i\}$ converge to some α_∞. This is followed by bounding the difference in value functions between an optimal policy (which need not lie in $\mathcal{P}^K$) and a policy that lies in $\mathcal{P}^K$ for sufficiently large K. This justifies the mean field model used in the previous section and our proof of convergence.

Proposition 1. *If $\{R_i\}$ is a decreasing sequence in i, then sequence $\{u_i\}$ is a decreasing sequence in i. Consequently, the sequence converges to some u_∞ in the limit as $i \to \infty$.*

Proof: Let the optimal policy for a device in state i be α_i^* for every i in $\{0, 1, ...\}$ and denote by u_i^* the value functions of the optimal policy. We define function $u_i(\alpha_i)$ as

$$u_i(\alpha_i) = \frac{\alpha_i(1-\gamma)}{1/\delta+\lambda+A-\alpha_i\gamma\beta}\big(R_i+\beta u_{-2}^*\big) - c(\alpha_i)\frac{1/\delta+\lambda+\alpha_i}{1/\delta+\lambda+A-\alpha_i\gamma\beta} + \beta\big(\frac{1/\delta}{1/\delta+\lambda+A-\alpha_i\gamma\beta}u_{i+1}^*\big) + \beta\big(\frac{\lambda}{1/\delta+\lambda+A-\alpha_i\gamma\beta}u_0^*\big).$$

From this definition, we have

$$u_i(\alpha_i^*) = u_i^* = \max_{\alpha} u_i(\alpha),$$

which implies that

$$\begin{aligned}u_i^* &\geq u_i(\alpha_{i+1}^*)\\ &= \frac{\alpha_{i+1}^*(1-\gamma)}{1/\delta+\lambda+\alpha_{i+1}^*(1-\gamma\beta)}\left(R_i+\beta u_{-2}^*\right)\\ &- c(\alpha_{i+1}^*)\frac{1/\delta+\lambda+\alpha_{i+1}^*}{1/\delta+\lambda+\alpha_{i+1}^*(1-\gamma\beta)}\\ &+ \beta\Big(\frac{1/\delta}{1/\delta+\lambda+\alpha_{i+1}^*(1-\gamma\beta)}u_{i+1}^*\Big)\\ &+ \beta\Big(\frac{\lambda}{1/\delta+\lambda+\alpha_{i+1}^*(1-\gamma\beta)}u_0^*\Big)\end{aligned}.$$

Note that we add the superscript $*$ to value function u_i to differentiate the notation from function $u_i(\alpha)$. u_i^* in this proof is the same as u_i defined in (3) and the statement of the proposition.

Note there is a similarity between $u_i(\alpha_{i+1}^*)$ and u_{i+1}^*. Namely,

$$\begin{aligned}u_i^* &\geq u_i(\alpha_{i+1}^*)\\ &= u_{i+1}^*+\beta\Big(\frac{1/\delta}{1/\delta+\lambda+\alpha_{i+1}^*(1-\gamma\beta)}\Big)(u_{i+1}^*-u_{i+2}^*)\end{aligned}.$$

Rearranging the terms, we get

$$\begin{aligned}&u_{i+1}^*-u_i^*\\ \leq&\beta\left(\frac{1/\delta}{1/\delta+\lambda+\alpha_{i+1}^*(1-\gamma\beta)}\right)(u_{i+2}^*-u_{i+1}^*)\\ \leq&\beta\Big(\frac{1/\delta}{1/\delta+\lambda}\Big)(u_{i+2}^*-u_{i+1}^*)\end{aligned}$$

for all i in $\{0,1,....\}$. This yields

$$\left(\beta\Big(\frac{1/\delta}{1/\delta+\lambda}\Big)\right)^J(u_{i+J+1}^*-u_{i+J+2}^*)\geq u_{i+1}^*-u_i^*.$$

Since the LHS of the inequality above converges to zero as $J\to\infty$, we have $u_i^*\geq u_{i+1}^*$. So, u_i^* is a decreasing sequence which is bounded above and below. Therefore, u_i^* converges to a fixed value. Let $u_\infty^* := \lim_{i\to\infty} u_i$. Then by definition, u_i^* converges to u_∞^*. ∎

Next, we state that α_i^* is a Cauchy sequence that converges to some α_∞.

Lemma 1. *If, R_i is a decreasing sequence in i, then the probing rate under the optimal policy, denoted by α_i^*, is also a decreasing sequence which converges to some α_∞^*.*

The proof follows from Proposition 1. Note that,

$$u_i^*-u_{i+1}(\alpha_i^*)\geq u_i^*-u_{i+1}^*\geq u_i(\alpha_{i+1}^*)-u_{i+1}^*$$

This inequality can be used to show α_i^* is a monotonic sequence in i. The rest of the proof is omitted due to space constraints.

Remark 2.

1) Both results generalize to an M-player game where the reward and costs are different for different players.
2) It is worth noting that Lemma (1) tells us that a device should probe most aggressively early on and look to drop the packet as the delay increases. Effectively, this observation seems to suggest at steady state the devices will behave such that the packets with the least delay will be more likely to be transmitted before packets with higher delay value.

The next result bounds the difference between a device that uses the optimal policy $\alpha^*=\{\alpha_0^*,\alpha_1^*,...\alpha_K^*,\alpha_{K+1}^*,...\}$ and the truncated version $\alpha^{(K)}:=\{\alpha_0^*,\alpha_1^*,...\alpha_K^*,0,0,.....\}$, assuming that all other devices choose the same fixed policy $\alpha^{(K)}$. We show that when γ is fixed, given any $\epsilon>0$, a device can chooses $\alpha^{(K)}$ for a sufficiently large K so that the difference between the two policies is at most ϵ, where K is independent of the number of devices. Therefore, a device may effectively choose a finite dimensional policy if it wishes to myopically optimize its utility function. This justifies our use of a finite dimensional policy while considering the mean field model to approximate the system.

Proposition 2. *Let α^* and $\alpha^{(K)}$ be as defined above. Given any $\epsilon>0$, there exists constant β_0 such that for all $\beta<\beta_0$ there exists K large enough so that*

$$|E_X\{u(X,\alpha^*)\}-E_{\tilde{X}}\{u(\tilde{X},\alpha^K)\}|<\epsilon,$$

where the expectation is taken over the stationary distribution of the device for fixed γ.

The proof of this proposition can be found in Appendix A. We have now characterized the mapping from γ to α. Given a fixed γ and a parameter K that is common among all other devices, a device may now use value iteration while fixing $\alpha_k=0$ for all $k>K$ to arrive at an approximate policy, which is the mapping T_2.

VI. Existence of MFNE

Based on the results in the previous sections, we will now show that there exists an MFNE using Brouwer's fixed point theorem. We show that both mappings T_1 and T_2 are Lipschitz. Therefore, there exists at least one fixed point (MFNE). This is stated in the theorem below.

Theorem 2. *There exists a constant β_0 such that for any $\beta<\beta_0$, a fixed point for the composite map $T_2\circ T_1$, denoted by T, exists.*

Proof: We will begin by showing that T_1 is Lipschitz. This turns out to be significantly easier than showing that T_2 is Lipschitz.

Claim: The map T_1 is Lipschitz in α.

Proof: To characterize T_1 as follows, given a policy, $\{x_i\}$ the detail balance equations yield (derived from Fig 1) give :

$$\pi_0=\frac{\lambda}{\lambda+1/\delta+x_0(1-\gamma)}$$

$$\pi_j=\pi_{j-1}\frac{1/\delta}{\lambda+1/\delta+x_j(1-\gamma)}$$

for $j>0$.

$$\pi_{-1}=\frac{1}{\lambda(1+\lambda)}\pi_{-2}$$

This gives us the following equation for the fraction of time spent in state -2 by a device,

$$1 = \pi_{-2}(1 + \frac{1}{\lambda(1+\lambda)}) + (1-\pi_{-2})\frac{\lambda}{1/\delta+\lambda+x_0(1-\gamma)} \times \sum_{i=0}^{\infty}\Pi_{j=1}^{i}\frac{1/\delta}{1/\delta+\lambda+x_j(1-\gamma)}.$$

.

Define,

$$\kappa := 1 + \frac{1}{\lambda(1+\lambda)} \tag{4}$$

and

$$\theta(x) := \frac{\lambda}{1/\delta+\lambda+x_0(1-\gamma)}\sum_{i=0}^{\infty}\Pi_{j=1}^{i}\frac{1/\delta}{1/\delta+\lambda+x_j(1-\gamma)} \tag{5}$$

the fraction of time a device spends in state -2 is given by :

$$\pi_{-2}(x) = \frac{1-\theta(x)}{\kappa-\theta(x)} \tag{6}$$

. Now, it can be checked that this function is differentiable and since,

$$\left|\frac{\partial\pi_{-2}}{\partial\alpha_i}\right| = \left|\frac{\kappa+1-2\theta(\alpha)}{(\kappa-\theta(\alpha))^2}\right|\left|\frac{\partial\theta(\alpha)}{\partial\alpha_i}\right|.$$

we need only check that $\|\frac{\partial\theta(\alpha)}{\partial\alpha}\|_1$ is finite to see that the map T_1 is in fact Lipschitz. ■

Checking that T_2 is continuous turns out to be more complicated. α_i depends on the entire sequence of u_i and u_i is both an implicit and explicit function of γ. Given that $\alpha \in \mathcal{P}^K$, we must first verify that $\{u_i\}$ is Lipschitz in γ with Lipschitz constant denoted by C_1.

Claim: If $\beta < \beta_0$, then u_i is Lipschitz in γ for all i. We provide a sketch of the proof for this claim due to space constraints.

Proof: Let $u_i^*(\gamma)$ be the optimal value function for fraction of busy channels γ. Then we can define $u(\alpha_i,\gamma)_i$ as we did in Proposition 1 for policy α_i. Now, if $\alpha_i^*(\gamma_1)$ and $\alpha_i^*(\gamma_2)$ are the optimal values for γ_1 and γ_2 respectively, then let,

$$\triangle_{i,\gamma_1,\gamma_2} := |u_i(\alpha_i^*(\gamma_1),\gamma_1) - u_i(\alpha_i^*(\gamma_2),\gamma_2)|$$

clearly,

$$\triangle_{i,\gamma_1,\gamma_2} \le |u_i(\alpha_i^*(\gamma_1),\gamma_1) - u_i(\alpha_i^*(\gamma_1),\gamma_2)|$$

If $\phi_0 = \frac{1}{A+\lambda+1/\delta-\alpha\gamma_1\beta}\left((1-\gamma_1)\alpha\beta^2\eta + \beta\lambda\right)$ and let $\phi_1 = \frac{\beta 1/\delta}{A+\lambda+1/\delta-\alpha\gamma_1\beta}$ one can use the inequality above to check that,

$$\triangle_{i,\gamma_1,\gamma_2} \le C_0|\gamma_1-\gamma_2| + \phi_0\triangle_{0,\gamma_1,\gamma_2} + \phi_1 C_1|\gamma_1-\gamma_2| + \phi_0\phi_1\triangle_{0,\gamma_1,\gamma_2} + \phi_1^2\triangle_{i+2,\gamma_1,\gamma_2}$$

One can expand the inequality for all i to bound $\triangle_{i,\gamma_1,\gamma_2}$. In particular this bound holds for $\triangle_{0,\gamma_1,\gamma_2}$:

$$\triangle_{0,\gamma_1,\gamma_2} \le \frac{C_0}{1-\phi_1}|\gamma_1-\gamma_2| + \frac{\phi_0}{1-\phi_1}\triangle_{0,\gamma_1,\gamma_2}$$

it follows that,

$$\triangle_{0,\gamma_1,\gamma_2} \le \frac{C_0}{1-\phi_0-\phi_1}|\gamma_1-\gamma_2|$$

Note that for $\beta < \beta_0$, $\phi_0 + \phi_1 < 1$ for all γ_1 and γ_2. Thus, for any γ_1 and γ_2, $\frac{\triangle_{0,\gamma_1,\gamma_2}}{\gamma_2-\gamma_1}$ is bounded. The rest of the proof then follows naturally. We can denote the Lipschitz constant by some C_1 ■

If $u_i^*(\gamma_1)$ and $u_i^*(\gamma_2)$ are the optimal value functions for γ_1 and γ_2 respectively, and if $|\gamma_1-\gamma_2| < \epsilon$ then it can be shown that

$$|u_i(\alpha_i^*(\gamma_1),\gamma_1) - u_i(\alpha_i^*(\gamma_2),\gamma_2)| < C_1\epsilon$$

by using first order Taylor expansion of $u_i(\gamma_2)$ about γ_1 and the triangle inequality,

$$\Bigg|\left[\frac{1}{A+\lambda+1/\delta-\alpha_i^*(\gamma_1)\beta\gamma_1} - \frac{1}{A+\lambda+1/\delta-\alpha_i^*(\gamma_2)\beta\gamma_2}\right]\times \Bigg(\alpha_i^*(\gamma_1)(1-\gamma_1)(R_i+\beta u_{-2}^*(\gamma_1)) - c(\alpha_i^*(\gamma_1))(\alpha_i^*(\gamma_1)+ \lambda+1/\delta) + \beta\lambda u_0^* + \beta 1/\delta u_{i+1}^*\Bigg) + \frac{1}{A+\lambda+1/\delta-\alpha_i^*(\gamma_1)\beta\gamma_1}\times \Bigg((\alpha_i^*(\gamma_2)-\alpha_i^*(\gamma_1))(1-\gamma_2)(R_i+\beta u_{-2}^*(\gamma_2)+ c(\alpha_i^*(\gamma_1))(\alpha_i^*(\gamma_1)+\lambda+1/\delta) - c(\alpha_i^*(\gamma_2))(\alpha_i^*(\gamma_2)+\lambda+1/\delta)\Bigg)\Bigg| < (C_1+C_2)\epsilon$$

for appropriate constants C_1 and C_2. It follows that $\sup_i|\alpha_i^*(\gamma_1)-\alpha_i^*(\gamma_2)| \xrightarrow{\gamma^{(1)}\to\gamma^{(2)}} 0$. From Proposition (2), there are only finitely many non zero α_i, this means that $\sum_{i=0}^{\infty}|\alpha_i^*(\gamma_1)-\alpha_i^*(\gamma_2)| \le K\sup_i|\alpha_i^*(\gamma_1)-\alpha_i^*(\gamma_2)|$ which converges to zero as $\gamma^{(1)}$ converges to $\gamma^{(2)}$. Thus, α is continuous in γ under the L1 norm. Therefore, the map T_2 is continuous in γ.

Therefore, the map $T:\alpha\to\alpha$ is also continuous in γ. The policy space, $\mathcal{P}^K$ is clearly convex. Since the domain of γ is compact and T_2 is continuous, so the range of T_2 must be compact. Therefore, by Brouwer's fixed point theorem, there exists a fixed value α^{**} such that

$$T(\alpha^{**}) = \alpha^{**}.$$

■

Having demonstrated that there exists a fixed point, our next theorem shows that the fixed point that is obtained using the map $T_2 \circ T_1$ is in fact an ϵ-Nash Equilibrium where ϵ converges to 0 as M and K tend to infinity.

Theorem 3. *The fixed point given by Theorem 2 is an ϵ-Nash equilibrium when the set of available policies are from $\mathcal{P}^K$, with $\epsilon \to 0$ as M tends to infinity.*

Proof: The main idea of the proof follows from the fact if one player chooses to deviate by choosing any policy in $\mathcal{P}^K$, then the mean field, γ changes by at most ϵ where ϵ goes to zero as M tends to infinity. This follows from a simple extension of the proof of ϵ-Nash Equilibrium in [1]. The proof

relies on Stein's method for finite state Markov chains.
Now, since the mean field γ only deviates by a small amount and $T_2 : \gamma \to \alpha$ is continuous, the best response policy in $\mathcal{P}^K$ will lie close to the mean field policy. ■

Here we would like to make few comments.

Remark 3.

1) Theorem 3 does not rely on asymptotic independence of the devices in our system. Therefore, it is nontrivial to show that the mean field remains unchanged when a finite set of players deviate. In fact under the typical law of large numbers would suggest $\epsilon = O(\sqrt{M})$, but in our case ϵ is O($M^{\frac{1}{3}}$).
2) Since $\mathcal{P}^K$ is ϵ close to an optimal policy for fixed γ, one might be tempted to state that the result of Theorem 3 holds for any policy instead of policies in $\mathcal{P}^K$. This is not straightforward since if a single node deviates with an arbitrary policy, the corresponding fraction of busy channels need not deviate by ϵ. Therefore, the best response need not necessarily be the MFNE policy.
3) While our theorems do not limit the number of fixed points, we strongly believe that there is a unique fixed point, primarily because we believe that the function T is decreasing in γ. Since our fixed points are the set of all γ such that $\gamma = T(\gamma)$, this would ensure that the fixed point is unique. This conjecture of a unique MFNE is a topic to be investigated later.

VII. Algorithm Design

The previous section proved that there exists at least one fixed point and the fixed point obtained is a local ϵ Nash Equilibrium. However, in the absence of contraction maps it is difficult to imagine how the device may achieve these equilibria. Here, we propose a scheme by which a device may achieve this equilibria.
We first note that while α can in general be a complicated variable even when it belongs to $\mathcal{P}^K$, the variable γ is a one dimensional real variable restricted to a closed bounded set, $(0, 1)$. For a fixed γ, a device may use policy iteration to find a policy α in $\mathcal{P}^K$ that is ϵ close to the optimal policy and for this policy α, explicitly compute $T_2(\alpha)$. This gives the device an estimated value of $T(\gamma)$ for a fixed value of γ. The device can now repeat this process for n such values of γ, and use this to estimate the function T computationally. Now, the device can find fixed points by solving $\gamma = T(\gamma)$. The device now picks the fixed point with the highest expected utility with ties broken based on lower fraction of busy channels.

Algorithm 1 Algorithm to find fixed points

input : Rewards $(R_j)_j$, $c(.)$, β, r_0, η, λ, δ
initialization : pick γ_i uniformly from the interval $(0, 1)$
while $i < n$ **do**
  $\alpha \leftarrow$ policy iteration(γ, $(R_j)_j$, $c(.)$, β, r_0, η, δ, λ)
  $T(i) \leftarrow \frac{1-\theta(\alpha)}{\kappa-\theta(\alpha)}$
  i ← i + 1
end while
interpolate T ;
find γ such that $\gamma = T(\gamma)$
output : γ

Ostensibly, one can view this algorithm as the device playing the game with itself in its own head to estimate the MFNE. Under these conditions, the MFNE assumptions of infinite players are justified and consistent thus, leading to local Nash equilibria for the devices. Since the devices are homogeneous, they will pick the same policies achieving the computed γ. An example implementation can be found in Figure 2.

VIII. Simulation Results

We present simulation results to demonstrate the performance of the proposed algorithm. We consider the following setting: $M = N$, i.e. the number of channels is equal to the number of devices, $K = 25$, $R_i = 10 \times 2^{-i}$, $\beta = 0.1$, $\delta = 1$, $A = 5$, and $c(\alpha) = 10\alpha^2$. We evaluate the delay experienced per packet when the system reaches MFNE and compare it to the throughput-oriented MAC protocol proposed in [1]. The protocols are:

- AD-MAC : This is our age-dependent distributed MAC. We varied the arrival rate λ from 0 to 2.
- D-MAC : This is the distributed MAC protocol proposed in [1]. We chose c to be 10 and evaluated the delay over the same range of λ. Our choices of c, λ in this case ensure that the MFNE exists as required in [1].

The per packet delays are shown in Figure 3, where we can observe that AD-MAC has significant smaller per packet delay when the arrival rate is low, and the delay of AD-MAC is always smaller than that of D-MAC in our simulations. In addition to the delay, we also compared the fraction of occupied channels, which reflects the system throughput. As we can observe from Figure 4, AD-MAC achieves higher throughput when $\lambda \leq 1.2$ (with smaller per packet delay as well). For $\lambda > 1.2$, the throughput is lower than that under D-MAC. This loss is to achieve lower per-packet delay.

IX. Conclusion

This paper studied a multichannel ultra-dense wireless network, and proposed an age-dependent distributed MAC where each device varies its idle-channel probing rate based on the age of the message it wants to transmit. The system has been analyzed using a mean field game framework. We characterized the policies that would be obtained through value iteration, and proved that there exist finite dimensional

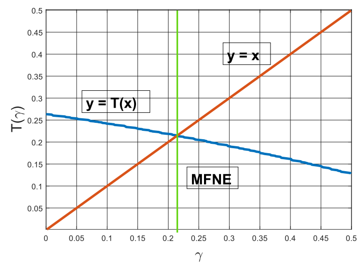

Figure 2. An example implementation with $K = 25$, $R_i = 2^{-i}$, $\beta = 0.1$, $\delta = 1$, $\lambda = 0.5$ and $c(\alpha) = 0.3\alpha^2$

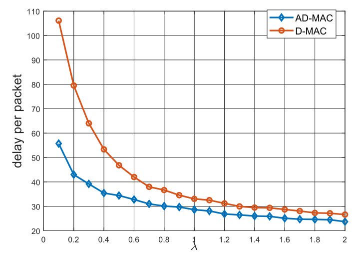

Figure 3. A comparison of the delays experienced per packet delivered over some time duration.

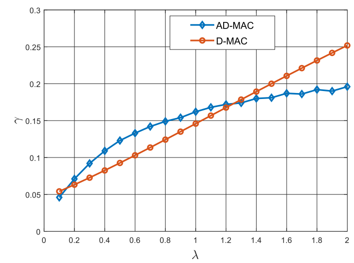

Figure 4. Comparing γ as a function of λ for the two protocols.

approximations with an additional condition on the discount factor. We provided an analysis of this protocol including the existence of fixed points in the mean field game formulation, which are ϵ-Nash equilibria in the original systems, with general convex cost functions and age-dependent rewards. Finally, we presented an algorithmic implementation of such a policy in the finite device setting for achieving the fixed points, and simulation results that showed the proposed algorithm achieved smaller per packet delay than earlier approaches.

Acknowledgement: This research was supported by grants NSF-Intel CNS-1719384, NSF CNS-1955696, CNS-2002608, CNS-2001687, ECCS-1609202, ECCS-1839816, CPS-2038963, ARO W911NF-19-1-0367 and ARO W911NF-19-2-024. DN would like to express gratitude towards JN and Dorabji Tata trusts for their support.

Appendix A
Policy Characterization

Proof of Proposition 2

Fix $\tilde{\epsilon} > 0$. Let $u_i^* = u_i(\alpha^*)$ as in the notation above and let $\tilde{u}_i = u_i(\alpha^K)$. We will begin by showing $u_0^* - \tilde{u}_0 < C_u\tilde{\epsilon}$ for some constant C_u for sufficiently large K when $\beta < \beta_0$.

$$\begin{aligned}
u_0^* - \tilde{u}_0 =& \Bigg(\frac{\beta\lambda}{\lambda + 1/\delta + \alpha_0(1-\gamma\beta)}(u_0^* - \tilde{u}_0) \\
&+ \eta\frac{\alpha_0(1-\gamma)}{\lambda + 1/\delta + \alpha_0(1-\gamma\beta)}(u_0^* - \tilde{u}_0)\Bigg) \\
&+ \frac{\beta 1/\delta}{\lambda + 1/\delta + \alpha_0(1-\gamma\beta)}(u_1^* - \tilde{u}_1^*) \\
&< \beta\left(\frac{\lambda}{\lambda + 1/\delta} + \eta\right)(u_0^* - \tilde{u}_{-1}) \\
&+ \frac{\beta\lambda}{\lambda + 1/\delta}\left(\beta\left(\frac{\lambda}{\lambda+1/\delta} + \eta\right)(u_0^* - \tilde{u}_0)\right. \\
&\left. + \frac{\beta\lambda}{\lambda + 1/\delta}(u_2^* - \tilde{u}_2)\right)
\end{aligned}$$

Expanding $u_j^* - \tilde{u}_j$ for $j > 0$ to $j < K$ we get:

$$\begin{aligned}
u_0^* - \tilde{u}_0 <& \beta\left(\frac{\lambda}{\lambda+1/\delta} + \eta\right)(u_0^* - \tilde{u}_0)\sum_{i=0}^{K}\left(\frac{\beta 1/\delta}{\lambda + 1/\delta}\right)^i \\
&+ \left(\frac{\beta 1/\delta}{\lambda+1/\delta}\right)^K(u_{K+1}^* - \tilde{u}_K)
\end{aligned}$$

For sufficiently large K, we can ensure that:

$$\begin{aligned}
u_0^* - \tilde{u}_0 <& \beta\left(\frac{\lambda}{\lambda+1/\delta} + \eta\right)\frac{1}{1 - \beta\frac{1/\delta}{1/\delta+\lambda}}(u_0^* - \tilde{u}_0) + \tilde{\epsilon} \\
<& \beta\left(1 + \eta\frac{\lambda + 1/\delta}{\lambda}\right)(u_0^* - \tilde{u}_0) + \tilde{\epsilon}
\end{aligned}$$

If we fix β_0 to be such that $\beta_0\big(1 + \eta\frac{\lambda+1/\delta}{\lambda}\big) < 1$, then $(u_0^* - \tilde{u}_0) < C_u\tilde{\epsilon}$. We can use the same reasoning to show that $u_i^* - \tilde{u}_i < C_u\tilde{\epsilon}$ for $0 < i < K/2$. We have shown that for sufficiently large K we can ensure that $|u_i^* - \tilde{u}_i| < C_u\tilde{\epsilon}$ for $0 \le i < K/2$.

For any policy $x := \{x_0, x_1, x_{-2} \ldots..\}$ with fixed γ, recall the detail balance equation from Claim (1), Theorem 2, the fraction of time for a device spent in state -2 is,

$$\pi_{-2}(x) = \frac{1-\theta(x)}{\kappa - \theta(x)}$$

with κ and $\theta(x)$ given by equation (4) and (5), respectively. Clearly, there exists K large enough so that $|\pi_{-2}(\alpha) - \pi_{-2}(\alpha^{(K)})| < \epsilon$. Since π_i converges geometrically to zero we can ensure that

- $K\epsilon \to 0$ as K tends to infinity.
- $\sum_{i=K/2}^{\infty}\pi_i < \epsilon$ for both α^* and α^K.
- $|\pi_i(\alpha^*) - \pi_i(\alpha^K)| < \epsilon$ for all $0 < i < K/2$.
- As above $|u_i^* - \tilde{u}_i| < \epsilon$ for all $0 < i < K/2$.

$$\begin{aligned}
&E_X\{u(X,\alpha^*)\} - E_{\tilde{X}}\{u(\tilde{X},\alpha^{(K)})\} \\
&= |\sum_{i=1}^{K/2}\left(\pi_i(\alpha)u_i^* - \pi_i(\alpha^K)\tilde{u}_i\right) \\
&+ \pi_0(\alpha^*)u_0^* - \pi_0(\alpha^K)\tilde{u}_0 + \pi_{-2}(\alpha^*)u_{-2} - \pi_{-2}(\alpha^K)\tilde{u}_{-2} \\
&+ \sum_{i=K/2+1}^{\infty}\left(\pi_i(\alpha^*)u_i^* - \pi_i(\alpha^K)\tilde{u}_i\right)|
\end{aligned}$$

Bounding $u_i^*(x)$ by $\frac{R}{1-\beta}$, replacing $\pi_i(\alpha^*)u_i^* - \pi_i(\alpha^K)\tilde{u}_i$ with $\pi_i(\alpha^K)(u_i^* - \tilde{u}_i) + u_i^*(\pi_i(\alpha^*) - \pi_i(\alpha^K))$, using the triangle inequality, for any $\hat{\epsilon} > 0$,

$$E_X\{u(X,\alpha^*)\} - E_{\tilde{X}}\{u(\tilde{X},\alpha^{(K)})\} < \hat{\epsilon}$$

References

[1] D. Narasimha, S. Shakkottai, and L. Ying, "A mean field game analysis of distributed MAC in ultra-dense multichannel wireless networks," *IEEE/ACM Transactions on Networking*, pp. 1–14, 2020.

[2] I. Hou, V. Borkar, and P. Kumar, "A theory of QoS for wireless," in *Proceedings of IEEE INFOCOM*, 2009.

[3] J. Jaramillo and R. Srikant, "Optimal scheduling for fair resource allocation in ad hoc networks with elastic and inelastic traffic," in *Proc. IEEE International Conference on Computer Communications (INFOCOM)*, San Diego, CA, March 2010.

[4] I. Hou and P. R. Kumar, "Scheduling heterogeneous real-time traffic over fading wireless channels," in *Proc. IEEE International Conference on Computer Communications (INFOCOM)*, San Diego, CA, March 2010.

[5] B. Li and A. Eryilmaz, "Optimal distributed scheduling under time-varying conditions: A Fast-CSMA algorithm with applications," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3278–3288, 2013.

[6] X. Kang, W. Wang, J. J. Jaramillo, and L. Ying, "On the performance of largest-deficit-first for scheduling real-time traffic in wireless networks," *IEEE/ACM Transactions on Networking*, vol. 24, no. 1, pp. 72–84, 2014.

[7] R. Singh and P. Kumar, "Throughput optimal decentralized scheduling of multihop networks with end-to-end deadline constraints: Unreliable links," *IEEE Transactions on Automatic Control*, vol. 64, no. 1, pp. 127–142, 2019.

[8] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE Int. Conf. Computer Communications (INFOCOM)*, 2012, pp. 2731–2735.

[9] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1807–1827, 2018.

[10] C. Kam, S. Kompella, G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, "On the age of information with packet deadlines," *IEEE Transactions on Information Theory*, vol. 64, no. 9, pp. 6419–6428, 2018.

[11] Y. Sun, I. Kadota, R. Talak, and E. Modiano, "Age of information: A new metric for information freshness," *Synthesis Lectures on Communication Networks*, vol. 12, no. 2, pp. 1–224, 2019.

[12] G. Bianchi, "Ieee 802.11-saturation throughput analysis," *IEEE Commun. Lett.*, vol. 2, pp. 318–320, 1998.

[13] F. Cecchi, S. C. Borst, J. S. H. van Leeuwaarden, and P. A. Whiting, "CSMA networks in a many-sources regime: A mean-field approach," in *Proc. IEEE Int. Conf. Computer Communications (INFOCOM)*, April 2016.

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

[15] K. Iyer, R. Johari, and M. Sundararajan, "Mean field equilibria of dynamic auctions with learning," *Management Science*, vol. 60, no. 12, pp. 2949–2970, 2014.

[16] J. Li, R. Bhattacharyya, S. Paul, S. Shakkottai, and V. Subramanian, "Incentivizing sharing in realtime D2D streaming networks: A mean field game perspective," *IEEE/ACM Trans. Netw.*, 2016.

[17] B. Xia, S. Shakkottai, and V. Subramanian, "Small-scale markets for bilateral resource trading in the sharing economy," *Proc. IEEE Int. Conf. Computer Communications (INFOCOM)*, 2018.

[18] A.-S. Sznitman, "Topics in propagation of chaos," in *Ecole d'été de probabilités de Saint-Flour XIX—1989*, 1991, pp. 165–251.

[19] C. Graham and S. Méléard, "Chaos hypothesis for a system interacting through shared resources," *Probability Theory and Related Fields*, vol. 100, no. 2, pp. 157–174, 1994.

[20] C. Bordenave, D. McDonald, and A. Proutiere, "Performance of random medium access control, an asymptotic approach," *SIGMETRICS Perform. Eval. Rev.*, vol. 36, no. 1, pp. 1–12, Jun. 2008.

[21] J.-w. Cho, J.-Y. Le Boudec, and Y. Jiang, "On the asymptotic validity of the decoupling assumption for analyzing 802.11 MAC protocol," *IEEE Trans. Inf. Theory*, vol. 58, June 2011.

[22] L. Ying, "On the approximation error of mean-field models," in *Proc. Ann. ACM SIGMETRICS Conf.*, Antibes Juan-les-Pins, France, June 2016.

[23] L. Ying, "Stein's method for mean field approximations in light and heavy traffic regimes," *Proc. ACM Meas. Anal. Comput. Syst.*, vol. 1, no. 1, pp. 12:1–12:27, Jun. 2017.

[24] N. Gast, "Expected values estimated via mean-field approximation are 1/n-accurate," *Proc. ACM Meas. Anal. Comput. Syst.*, vol. 1, no. 1, pp. 17:1–17:26, Jun. 2017.

[25] N. Gast and B. Van Houdt, "A refined mean field approximation," in *Proc. Ann. ACM SIGMETRICS Conf.*, Irvine, CA, 2018.