A Mean Field Game Approach to Scheduling in Cellular Systems

Mayank Manjrekar*, Vinod Ramaswamy*, and Srinivas Shakkottai*
*Dept. of ECE, Texas A&M University, Email: {mayankm, vinod83, sshakkot}@tamu.edu

***Abstract*—We study auction-theoretic scheduling in cellular networks using the idea of mean field equilibrium (MFE). Here, agents model their opponents through a distribution over their action spaces and play the best response. The system is at an MFE if this action is itself a sample drawn from the assumed distribution. In our setting, the agents are smart phone apps that generate service requests, experience waiting costs, and bid for service from base stations. We show that if we conduct a second-price auction at each base station, there exists an MFE that would schedule the app with the longest queue at each time. The result suggests that auctions can attain the same desirable results as queue-length-based scheduling. We present results on the asymptotic convergence of a system with a finite number of agents to the mean field case, and conclude with simulation results illustrating the simplicity of computation of the MFE.**

I. Introduction

There has recently been a rapid increase in the use of smart hand-held devices for Internet access. These devices are supported by cellular data networks, which carry the packets generated by apps running on these smart devices. We consider a system consisting of smart phone users whose apps are modeled as queues that arrive when the user starts the app, and depart when the user terminates that app. Packets generated by apps are captured as jobs that arrive to the queues. Users move around cells that each has a base station, and scheduling a particular user provides service to the queue that represents his/her currently running app. User interest could shift from app to app, regardless of whether or not there are buffered packets. Hence, a queue might terminate and be replaced by a new one even if there are jobs waiting for service. The problem that we consider is that of fair scheduling in such a system of many base stations and ephemeral queues.

Most work on scheduling has focused on the case of a finite number of infinitely long lived flows, with the objective being to maximize the total throughput. A seminal piece of work under this paradigm introduced the so-called *max-weight algorithm* [1]. Here, the drift of a quadratic Lyapunov function is minimized by maximizing the queue-length weighted sum of acceptable schedules. Later work (*e.g.*, [2]–[4]) has used a similar approach in a variety of network scenarios. If queues arrive and depart, then a natural scheduling policy in the single server case is a *Longest-Queue-First (LQF)* scheme, in which each server serves the longest of the queues requesting service from it. LQF has many attractive properties, such a minimizing the expected value of the longest queue in the system.

Critical to the above approach is the assumption that the queue length values are given to the scheduler. While the downlink queue lengths would naturally be available at a cellular base station, the only way to obtain uplink queue information is to ask the users themselves. However, a larger value of queue length results in a higher probability of being scheduled under all the above policies, implying an incentive to lie. How are we to design a scheduling scheme that possesses the good qualities of LQF, while relying on self-reported values?

An appealing idea is to use a pricing scheme to inform scheduling decisions for cellular data access. For instance, [5] describes an experimental trial of a system in which day-ahead prices are announced to users, who then decide on whether or not to use their 3G service based on the price at that time. However, these prices have to be determined through trial and error. Can we determine prices by using an auction?

Our key objective is to design an incentive compatible scheduling scheme that behaves in an LQF-like fashion. We consider a system in which each app bids for service from the base station that it is currently connected to. The auction is conducted in a second-price fashion, with the winner being the app that bids highest, and the charge being the value of the second highest bid. It is well known that such an auction promotes truth-telling [6]. Would the scheduling decisions resulting from such auctions resemble that of LQF? Would conducting such an auction repeatedly over time with queues arriving and departing result in some form of equilibrium?

Mean Field Games

We investigate the existence of an equilibrium using the theory of Mean Field Games (MFG) [7]–[12]. In MFG, the players assume that each opponent would play an action drawn *independently* from a fixed distribution over its action space. The player then chooses an action that is the best response against actions drawn in this fashion. We say that the system is at Mean Field Equilibrium (MFE) if this best response action is itself a sample drawn from the assumed distribution.

The MFG framework offers a structure to approximate so-called Perfect Bayesian Equilibrium (PBE) in dynamic games. PBE requires each player to keep track of their beliefs on the future plays of every other opponent in the system, and play the best response to that belief. This makes the computation

Research was funded in part by NSF grants CNS-0904520, CNS-0963818, CNS-1149458, DTRA grant HDTRA 1-13-1-0030 and AFOSR FA9550-13-1-0008. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

978-1-4799-3360-0/14/$31.00 © 2014 IEEE

of PBE intractable when the number of players is large. The MFG framework simplifies computation of the best response, and often turns out to be asymptotically optimal.

Work on MFGs has mostly focused on showing the existence, accuracy and stability of MFE [7]–[10]. In the space of queueing systems, some recent work considers the game of sampling a number of queues and joining one [12]. However, ours is a scheduling problem in queueing system interacting with an auction system, which we believe is unique.

Our methodology is close to work by Iyer *et al.* [11], where the model is that of advertisers competing via a second price auction for spots on a webpage. The bid must lie in a finite real interval, and the winner can place an ad on the webpage. With time, the advertisers learn about the value of winning (probability of making a sale). Verifying the conditions of the Schauder Fixed Point Theorem shows the existence of the MFE, and the results of Graham *et al.* [13] are used to show that the MFG approximation is asymptotically valid. Our analytical approach is similar at a high level, but we have two main differences. The first is that we allow bids to lie in the full positive real line, and the second is that we have a queueing system with job arrivals. These differences require a full reworking of the proofs in our setting.

Overview of Model and Results

Our Bayesian game is conducted in discrete time, and consists of N cells and NM agents (apps), which are randomly permuted in these cells at each time instant, with each cell having exactly M agents. Each cell contains a base station, which conducts a second price auction to choose which agent to serve. Each agent must choose its bid in response to its state and its belief over the bids of its competitors.

Figure 1 illustrates the MFG approximation, with the auction (shown as blue/dark tiles) and the queue dynamics (shown as beige/light tiles). An MFG is described from the perspective of a single agent, which assumes that the actions of all its competitors are drawn independently from some distribution.

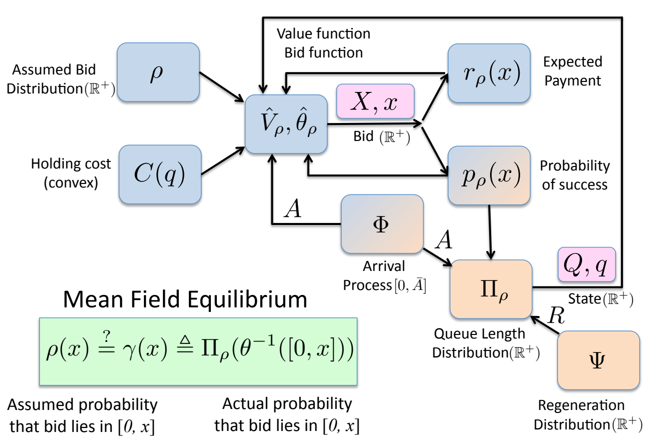

Fig. 1. The game consists of an auction part (blue/dark tiles) and a queue dynamics part (beige/light tiles). The system is at MFE if the distribution of the bid X is equal to the assumed bid distribution ρ.

Auction System: The agent of interest competes in a second price auction against $M - 1$ other agents, whose bids are assumed to be independently drawn from a continuous, finite mean (cumulative) bid distribution ρ, with support $\mathbb{R}^+$. The state of the agent is its current queue length q (the random variable is represented by Q). The queue length induces a holding cost $C(q)$, where $C(.)$ is a strictly convex and increasing function. Suppose that the agent bids an amount $x \in \mathbb{R}^+$. The outcomes of the auction are that the agent would obtain a unit of service with probability $p_\rho(x)$ and would have to pay an expected amount of $r_\rho(x)$ when all the other bids are drawn independently from ρ. Further, the queue has future job arrivals according to distribution Φ, with the random job size being denoted by A. Finally, the app can terminate at any time instant with probability $1 - \beta$. Based on these inputs, the agent needs to determine the value of its current state $\hat{V}_\rho(q)$, and the best response bid to make $x = \hat{\theta}_\rho(q)$.

Queueing System: The queueing dynamics are driven by the arrival process Φ and the probability of obtaining service being $p_\rho(x)$ as described above. When the user terminates an app, he/she immediately starts a fresh app, *i.e.,* a new queue takes the place of a departing queue. The initial condition of this new queue is drawn from a regeneration distribution Ψ, whose support is $\mathbb{R}^+$. The stationary distribution associated with this queueing system (if it exists) is denoted Π_ρ.

Mean Field Equilibrium: The probability that the agent's bid lies in the interval $[0, x]$ is equal to the probability that the agent's queue length lies in some set whose best response is to bid between $[0, x]$. Thus, the probability of the bid lying in the interval $[0, x]$ is $\Pi_\rho(\hat{\theta}_\rho^{-1}([0, x]))$, which we define as $\gamma(x)$. According to the assumed (cumulative) bid distribution, the probability of the same event is $\rho(x)$. If $\rho(x) = \gamma(x)$, it means that the assumed bid distribution is consistent with the best response bid distribution, and we have an MFE.

Organization and Main Results: After preliminaries on Bayesian and mean field systems in Sections II–III, we consider the problem of determining the cost minimizing bid function and the corresponding value function as a Markov Decision Process in Section IV. We show that the Bellman operator corresponding to the MDP is a contraction mapping with a unique minimum, implying that value iteration would converge to the best response bid. Further, we show that the best response bid is monotone increasing in queue length.

We next prove the existence of the MFE in Sections V–VI by verifying the conditions of the Schauder Fixed Point Theorem. We need to show that the mapping between ρ to γ is continuous, and that the space in which γ lies is contained in a compact subset of the space from which ρ is drawn. In order to do this, we first show that the mapping between ρ and $\hat{\theta}_\rho(.)$ is continuous, and then show that the map between ρ and $\Pi_\rho(.)$ is continuous. Putting these together yields the required continuity conditions. We then verify the conditions of the Arzelà-Ascoli Theorem for showing compactness.

We show in Section VII that the MFE in our case is an asymptotically accurate approximation of a PBE. The result follows from the fact that any finite subset of users is unlikely to have interacted with any of the others as N becomes large. Finally, we present simulation results in Section VIII, showing that MFE computation is straightforward.

Discussion: The bid function $\hat{\theta}_\rho(q)$ is monotone increasing in q regardless of ρ. This implies that the service regime corresponding to MFE is identical to LQF (or a weighted version if we have different user classes). Further, our simulations suggest that if the base stations were to compute the empirical bid distribution and return it to the users, the eventual bid distribution would be the MFE. Thus, the desirable properties of LQF are a natural result of auction-based scheduling.

II. Bayesian Game Model

We consider a system with N cells and NM agents. We assume that time is discrete, and at each time instant, the NM agents are randomly permuted among the N cells, with exactly M agents in each. Each cell then conducts a second price auction. The winner of an auction gets service at that instant; reducing the workload in the winner's queue by at most one unit. Let $Q_{i,k}$ represent the residual workload of agent i, just before k^{th} auction. We assume that $Q_{i,k} \in [0, \infty)$ and note that it completely represents the state of the queue at time k. Agent i's workload is influenced by the following:

1) *Arrivals:* After every auction, an arrival $A_{i,k}$ occurs at agent i, where $A_{i,k}$ is a random variable independent of every other parameter and distributed according to Φ.
2) *Service:* $D_{i,k}$ is the random variable representing the amount of service delivered at the k-th time instant. We assume that the server serves at-most a unit workload of the winner in any auction. $D_{i,k} = \min\{1, Q_{i,k}\} \times W_{i,k}$, where $W_{i,k} = \mathbf{1}(i \text{ wins at time } k)$.
3) *Regeneration:* We assume that agent i may regenerate with probability $1 - \beta$ after participating in an auction, where $0 < \beta < 1$. We assume that the new workload is a random variable distributed according to Ψ.

Hence, the state of agent i at time $k + 1$ is,

$$Q_{i,k+1} = \begin{cases} R_{i,k} & \text{if regenerates at } k \\ Q_{i,k} - D_{i,k} + A_{i,k} & \text{otherwise,} \end{cases} \tag{1}$$

where $A_{i,k} \sim \Phi$ and $R_{i,k} \sim \Psi$. Below we state the assumptions on the arrival and regeneration processes.

Assumption 1. *The arrivals $\{A_{i,k}\}$ are i.i.d random variables distributed according to Φ. We assume that $A_{i,k} \in [0, \bar{A}]$, where $\bar{A}$ is finite. Also, these random variables have a bounded density function, ϕ (i.e., $||\phi|| < c_\phi$, where $||.||$ is the sup norm).*

Assumption 2. *The regeneration values $\{R_{i,k}\}$ are i.i.d random variables distributed according to Ψ, and they have a bounded density ψ (i.e., $\|\psi\| < c_\psi$, where $||.||$ is the sup norm).*

Each agent bears a *holding* cost at every instant, that corresponds to the dis-utility due to unserved workload. The holding cost of agent i at time k is $C(Q_{i,k})$, where $C : \mathbb{R}^+ \to \mathbb{R}^+$. The agent also pays for service if it wins the auction. This is called the bidding cost. Let $X_{i,k}$ be the bid submitted by agent i in the k^{th} auction and

$$\bar{X}_{-i,k} = \max_{j \in M_{i,k}} X_{j,k},$$

where $M_{i,k}$ is the set consisting of all other agents participating in at time k with agent i. Then, the bidding cost of agent i is $\bar{X}_{-i,k} \times W_{i,k}$. We make a convexity assumption on the holding cost, which is motivated by delay usually being convex in queue length.

Assumption 3. *The holding cost function $C : \mathbb{R}^+ \mapsto \mathbb{R}^+$ is continuous, increasing and strictly convex. We also assume that C is $O(q^m)$ for some integer m.*

The polynomial form above is for technical reasons, but is not very restrictive since many convex functions can be approximated quite well.

Optimal bidding strategy

The information available with any agent about the market at any time prior to the auction only includes the following:

1) The bids it made in each previous auction from last regeneration.
2) The auctions that it won.
3) The payments made for the auctions won.

Let $H_{i,k}$ be the history vector containing the above information available to agent i at time k. Each agent holds a belief that is a distribution over future trajectories, which gets updated via Bayes' rule as new information arrives at the occurrence of each auction. Let $\mu_{i,k}$ be the belief of agent i at time k.

Let *pure strategy* θ_i be the history dependent strategy of agent i, i.e $\theta_i(H_{i,k}) = X_{i,k}$. We define $\boldsymbol{\theta}_{-i}$ to be the vector of strategies of all agents except agent i and $\boldsymbol{\theta} = [\theta_i, \boldsymbol{\theta}_{-i}]$. We refer to $\boldsymbol{\theta}$ as *strategy profile*.

Given a strategy profile $\boldsymbol{\theta}$, a history vector $H_{i,k}$ and a belief vector $\mu_{i,k}$, the expected cost is

$$V_{i,\mu_{i,k}}(H_{i,k}; \boldsymbol{\theta}) = \mathbb{E}_{\theta,\mu_{i,k}} \left[\sum_{t=k}^{T_i^k} \left[C(Q_{i,t}) + \bar{X}_{-i,t} \mathbf{1}(W_{i,t} = 1) \right] \right], \tag{2}$$

where T_i^k is the time at which i regenerates after time k.

We now introduce the notion of Nash equilibrium in dynamic games, called *Perfect Bayesian Equilibrium* (PBE).

Definition 1. *A strategy profile $\boldsymbol{\theta}$ is said to be a PBE if*

1) *For each agent i, after any history $H_{i,k}$, the strategy $\theta_i \in \arg\max_{\theta_i'} V_{i,\mu_{i,k}}(H_{i,k}, \theta_i', \boldsymbol{\theta}_{-i})$*
2) *The belief vectors $\mu_{i,k}$ are updated via Bayes' rule for all agents.*

The Bayesian game framework requires each agent to keep track of complex beliefs over other agents, which becomes computationally prohibitive in large systems. We next describe the mean field model.

III. Mean Field Model

As discussed in the Introduction, the mean field model is an approximation of the Bayesian game as the number of agents approaches infinity. Here, according to the belief of a single agent, the distribution of a random agent's state does not change under Bayesian updates. Further, the bid distributions of the other $M - 1$ agents in an auction are independent, as it is unlikely that they would have interacted from the point of earliest regeneration of any of the agents in the auction. The identity the other agents is irrelevant, and so agent i needs to only maintain belief over the bid of a random agent.

A. Agent's decision problem

Let the candidate be agent i. Suppose that the belief over the bid of a random agent has cumulative distribution ρ. We assume that $\rho \in \mathcal{P}$ where,

$$\mathcal{P} = \{G | G \text{ is a continuous c.d.f, } \int (1 - G(x))dx < E\},$$

for some $E < \infty$, to be defined later.

Since the time of regeneration T_i^k is a geometric random variable, the expected cost of the agent given in (2) can be can be re-written as

$$V_{i,\rho}(H_{i,k}; \theta_i) = \mathbb{E}\left[\sum_{t=k}^{\infty} \beta^t [C(Q_{i,t}) + r_\rho(X_{i,t})]\right], \quad (3)$$

where the expectation is over future state evolutions. By replacing the belief with ρ, we have made an agent's decision problem independent of other agents' strategies. Hence, we represent the cost by $V_{i,\rho}(H_{i,k}; \theta_i)$. Also, $r_\rho(x) = \mathbb{E}[\bar{X}_{-i,k}\mathbf{I}\{\bar{X}_{-i,k} \le x\}]$ is the expected payment when i bids x under the assumption that the bids of other agents are distributed according to ρ. Hence, given ρ, the win probability in the auction is

$$p_\rho(x) = \mathbb{P}(\bar{X}_{-i,k} \le x) = \rho(x)^{M-1}. \quad (4)$$

The expected payment when bidding x is

$$\begin{aligned} r_\rho(x) &= \mathbb{E}[\bar{X}_{-i,k}\mathbf{I}\{\bar{X}_{-i,k} \le x\}] \\ &= x p_\rho(x) - \int_0^x p_\rho(u)du. \end{aligned} \quad (5)$$

The state process $Q_{i,k}$ is Markov and has an invariant transition kernel

$$\begin{aligned} &\mathbb{P}(Q_{i,k+1} \in B | Q_{i,k} = q, X_{i,k} = x) = \\ &\quad \beta p_\rho(x)\mathbb{P}((q-1)^+ + A_k \in B) \\ &\quad + \beta(1 - p_\rho(x))\mathbb{P}(q + A_k \in B) + (1-\beta)\Psi(B), \end{aligned} \quad (6)$$

where $B \subseteq R^+$ is a Borel set and $x^+ \triangleq \max(x, 0)$. Recall that $A_k \sim \Phi$ is the arrival between $(k)^{th}$ and $(k+1)^{th}$ auction and Ψ is density function of the regeneration process. In the above expression, the first term corresponds to the event that agent wins the auction at time k, while the second corresponds to the event that it does not. The last term captures the event that the agent regenerates after auction k. The agent's decision problem can be modeled as an infinite horizon discounted cost MDP. It can be shown that there exists an optimal Markov deterministic policy for our MDP [14]. Then, from (3), the optimal value function of the agent can be written as

$$\begin{aligned} &\hat{V}_{i,\rho}(q) = \\ &\quad \inf_{\theta_i \in \Theta} \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^t [C(Q_{i,t}) + r_\rho(X_{i,t})] \,|Q_{i,0} = q\right], \end{aligned} \quad (7)$$

where Θ is the space of Markov deterministic policies.

Note that user index is redundant in the above expression as we are concerned with a single agent's decision problem, In future notations, we will omit the user subscript i.

B. Stationary distribution

Given cumulative bid distribution ρ and a Markov policy $\theta \in \Theta$, the transition kernel given by (6) can be re-written as,

$$\begin{aligned} &\mathbb{P}(Q_{k+1} \in B | Q_k = q) = \\ &\quad \beta p_\rho(\theta(q))\mathbb{P}((q-1)^+ + A_k \in B) \\ &\quad + \beta(1 - p_\rho(\theta(q)))\mathbb{P}(q + A_k \in B) + (1-\beta)\Psi(B). \end{aligned} \quad (8)$$

Then, we have an important result in the following lemma:

Lemma 1. *The Markov chain described by the transition probabilities in* (8) *is positive Harris recurrent and has a unique stationary distribution.*

Proof: From (8) we note that,

$$\mathbb{P}(Q_{k+1} \in B | Q_k = q) \ge (1-\beta)\Psi(B)$$

where $0 < \beta < 1$ and Ψ is a probability measure. The result then follows from results in Chapter 12, Meyn and Tweedie [15]. ■

We denote the unique stationary distribution by $\Pi_{\rho,\theta}$.

C. Mean field equilibrium

As described in the Introduction, the mean field equilibrium requires the consistency check that the bid distribution γ induced by the stationary distribution Π_{ρ,θ_ρ} should be equal to the bid distribution conjectured by the agent, i.e., ρ. Thus, we have the following definition of MFE:

Definition 2 (Mean field equilibrium)**.** *Let ρ be a bid distribution and θ_ρ be a stationary policy for an agent. Then, we say that (ρ, θ_ρ) constitutes a mean field equilibrium if*

1) *θ_ρ is an optimal policy of the decision problem in (7), given bid distribution ρ; and*
2) *$\rho(x) = \gamma(x) \triangleq \Pi_\rho(\theta_\rho^{-1}([0, x])), \forall x \in \mathbb{R}^+$, where $\Pi_\rho = \Pi_{\rho,\theta_\rho}$.*

A main result of this work is showing the existence of an MFE. Due to space limitations, we will only present proof sketches in most cases. All proofs can be found in our technical report [16].

IV. Properties of Optimal Bid Function

The decision problem given by (7) is an infinite horizon, discounted Markov decision problem. The optimality equation or Bellman equation corresponding to the decision problem is

$$\begin{aligned} \hat{V}_\rho(q) &= C(q) + \beta\mathbb{E}_A(\hat{V}_\rho(q+A)) + \inf_{x \in R^+} [r_\rho(x) \\ &-p_\rho(x)\beta\mathbb{E}_A\left(\hat{V}_\rho(q+A) - \hat{V}\rho((q-1)^+ + A)\right)\Big], \end{aligned} \quad (9)$$

where $A \sim \Phi$, and we use the notation $\max(0, z) = z^+$.

Define the set of functions

$$\mathcal{V} = \left\{f : \mathbb{R}^+ \mapsto \mathbb{R}^+ : \sup_{q \in \mathbb{R}^+} \left|\frac{f(q)}{w(q)}\right| < \infty\right\}, \quad (10)$$

where $w(q) = \max\{C(q), 1\}$. Note that $\mathcal{V}$ is a Banach space with *w-norm*,

$$\|f\|_w = \sup_{q \in \mathbb{R}^+} \left|\frac{f(q)}{w(q)}\right| < \infty.$$

Also, define the operator T_ρ as

$$(T_\rho f)(q) = C(q) + \beta \mathbb{E}_A f(q+A) + \inf_{x\in\mathbb{R}^+} \big[r_\rho(x) \\ -p_\rho(x)\beta(\mathbb{E}_A(f(q+A) - f((q-1)^+ + A)))\big], \quad (11)$$

where $f \in \mathcal{V}$. It is straightforward to show that the infimum in the above operator occurs at

$$\beta \Delta f(q)^+, \quad (12)$$

where $\Delta f(q) = \mathbb{E}_A(f(q+A) - f((q-1)^+ + A))$. Then, substituting from (4), (5) and (12), (11) can be rewritten as

$$(T_\rho f)(q) = \\ C(q) + \beta \mathbb{E}_A f(q+A) - \int_0^{\beta\Delta f(q)^+} p_\rho(u) du. \quad (13)$$

The following lemma characterizes the optimal solution.

Lemma 2. *Given a cumulative bid distribution ρ,*

1) *There exists a $j \in \mathbb{N}$ such that $T_\rho^j : \mathcal{V} \to \mathcal{V}$ is a contraction mapping. Hence, there exists a unique $f_\rho^* \in \mathcal{V}$ such that $T_\rho f_\rho^* = f_\rho^*$, and for any $f \in \mathcal{V}$, $T_\rho^n f \to f_\rho^*$ as $n \to \infty$.*
2) *The fixed point f_ρ^* of operator T_ρ is the unique solution to the optimality equation (9), i.e., $f_\rho^* = \hat{V}_\rho$.*
3) *Let,*
$$\hat{\theta}_\rho(q) = \beta \Delta \hat{V}_\rho(q)^+.$$
Then, $\hat{\theta}_\rho$ is an optimal policy.

Proof: We first show that $T_\rho f \in \mathcal{V}, \forall f \in \mathcal{V}$. The proof then follows through a verification of the conditions of Theorem 6.10.4 in [17]. ■

Corollary 3. *An optimal policy of the agent's decision problem* (7) *is given by*

$$\hat{\theta}_\rho(q) = \beta \mathbb{E}_A \left[\hat{V}_\rho(q+A) - \hat{V}_\rho((q-1)^+ + A)\right].$$

We now establish that $\hat{V}_\rho$ and $\hat{\theta}_\rho$ are continuous and increasing functions.

Lemma 4. *Given a cumulative bid distribution function ρ*

1) *$\hat{V}_\rho$ is a continuous increasing function.*
2) *$\hat{\theta}_\rho$ is a continuous strictly increasing function.*

Proof: Let $f \in \mathcal{V}$. Suppose f is a continuous monotone increasing function. We first prove that $T_\rho f$ is also continuous monotone increasing function. Since, $T_\rho^n f \to \hat{V}_\rho$ according to Statement 2 of Lemma 2, we conclude that $\hat{V}_\rho$ also has the same property.

Let $q > q'$. Then,

$$\begin{aligned}
&T_\rho f(q) - T_\rho f(q') = C(q) - C(q') \\
&+ \beta \mathbb{E}_A(f(q+A) - f(q'+A)) \\
&+ \inf_x [r_\rho(x) - \beta p_\rho(x) \mathbb{E}_A(f(q+A) - f((q-1)^+ + A))] \\
&- \inf_b [r_\rho(x) - \beta p_\rho(x) \mathbb{E}_A(f(q'+A) - f((q'-1)^+ + A))] \\
&\overset{(a)}{\geq} \beta \mathbb{E}_A(f(q+A) - f(q'+A)) \\
&\quad + \beta \inf_b \big[p_\rho(x) \mathbb{E}_A(f(q'+A) - f((q'-1)^+ + A) \\
&\qquad - f(q+A) + f((q-1)^+ + A))\big] \\
&\geq \beta \min \{\mathbb{E}_A(f(q+A) - f(q'+A)), \\
&\qquad \mathbb{E}_A(f((q-1)^+ + A) - f((q'-1)^+ + A))\} \overset{(b)}{\geq} 0,
\end{aligned}$$

where (a) follows from the assumption that $C(.)$ is an increasing function, and (b) follows from the assumption that $f(.)$ is an increasing function.

To prove that $T_\rho f$ is continuous consider a sequence $\{q_n\}$ such that $q_n \to q$. Since f is a continuous function, $f(q_n + a) \to f(q+a)$. Then, by using dominated convergence theorem, we have $\mathbb{E}_A f(q_n + A) \to \mathbb{E}_A f(q+A)$ and $\mathbb{E}_A f((q_n - 1)^+ + A) \to \mathbb{E}_A f((q-1)^+ + A)$. Also, $\Delta f(q_n) \geq 0$ as f is an increasing function. Then, from (13), we get that

$$T_\rho f(q_n) = C(q_n) + \beta \mathbb{E}_A f(q_n + A) - \int_0^{\beta \Delta f(q_n)} p_\rho(u) du \\ \to C(q) + \beta \mathbb{E}_A f(q+A) - \int_0^{\beta\Delta f(q)} p_\rho(u) du = T_\rho f(q).$$

Hence, $T_\rho f$ is a continuous function. This yields Statement 1 in the lemma.

Now, to prove the second part, assume that Δf is an increasing function. First, we show that $\Delta T_\rho f$ is an increasing function. Let $q > q'$. From (13), for any $a < \bar{A}$ we can write

$$\begin{aligned}
&(T_\rho f)(q+a) - (T_\rho f)((q-1)^+ + a) \\
&\qquad - (T_\rho f)(q'+a) + (T_\rho f)((q'-1)^+ + a) \\
=&C(q+a) - C((q-1)^+ + a) \\
&- C(q'+a) + C((q'-1)^+ + a) \\
&+ \beta \mathbb{E}_A f(q+a+A) - \beta \mathbb{E}_A f((q-1)^+ + a + A) \\
&- \beta \mathbb{E}_A f(q'+a+A) + \beta \mathbb{E}_A f((q'-1)^+ + a + A) \\
&- \int_{\beta\Delta f(q'+a)}^{\beta\Delta f(q+a)} p_\rho(u)\ du + \int_{\beta\Delta f((q'-1)^+ + a)}^{\beta\Delta f((q-1)^+ + a)} p_\rho(u)\ du \\
=&C(q+a) - C((q-1)^+ + a) \\
&- C(q'+a) + C((q'-1)^+ + a) \\
&+ \beta \mathbb{E}_A f((q+a-1)^+ + A) - \beta \mathbb{E}_A f((q-1)^+ + a + A) \\
&- \beta \mathbb{E}_A f((q'+a-1)^+ + A) + \beta \mathbb{E}_A f((q'-1)^+ + a + A) \\
&+ \int_{\beta\Delta f(q'+a)}^{\beta\Delta f(q+a)} 1 - p_\rho(u)\ du + \int_{\beta\Delta f((q'-1)^+ + a)}^{\beta\Delta f((q-1)^+ + a)} p_\rho(u)\ du
\end{aligned}$$

It can be easily verified that $\mathbb{E}_A(f(q+a-1)^+ + A) - \mathbb{E}_A(f(q-1)^+ + a + A) - \mathbb{E}_A(f(q'+a-1)^+ + A) + \mathbb{E}_A(f(q'-1)^+ + a + A) \geq 0$ as f is increasing (due to Statement 1 of this lemma). From the assumption that Δf is increasing, the last two terms in the above expression are also non-negative. Now, taking expectation on both sides, we obtain $\Delta T_\rho f(q) - \Delta T_\rho f(q') \geq \Delta C(q) - \Delta C(q') > 0$. Therefore, from Statements 2 and 3 of Lemma 2, we have

$$\hat{\theta}_\rho(q) - \hat{\theta}_\rho(q') = \Delta \hat{V}_\rho(q) - \Delta \hat{V}_\rho(q') \geq \Delta C(q) - \Delta C(q') > 0.$$

Here, the last inequality holds since C is a strictly convex increasing function. ■

V. Existence of MFE

The main result showing the existence of MFE is as follows.

Theorem 5. *There exists an MFE $(\rho, \hat{\theta}_\rho)$ such that*

$$\rho(x) = \gamma(x) \triangleq \Pi_\rho \left(\hat{\theta}_\rho^{-1}[0, x]\right), \forall x \in \mathbb{R}^+.$$

We first introduce some useful notation. Let $\Theta = \{\theta : \mathbb{R} \mapsto \mathbb{R}, \sup_{q \in \mathbb{R}^+} \left|\frac{\theta(q)}{w(q)}\right| < \infty\}$. Note that Θ is a normed space with w-norm. Also, let Ω be the space of absolutely continuous probability measures on $\mathbb{R}^+$. We endow this probability space with the topology of weak convergence. Note that this is same as the topology of point-wise convergence of continuous cumulative distribution functions.

We define $\theta^* : \mathcal{P} \mapsto \Theta$ as $(\theta^*(\rho))(q) = \hat{\theta}_\rho(q)$, where $\hat{\theta}_\rho(q)$ is the optimal bid given by Corollary 3. It can easily verified that $\hat{\theta}_\rho \in \Theta$. Also, define the mapping Π^* that takes a bid distribution ρ to the invariant workload distribution $\Pi_\rho(\cdot)$. Later, using Lemma 8 we will show that $\Pi_\rho(\cdot) \in \Omega$. Therefore, $\Pi^* : \mathcal{P} \to \Omega$. Finally, define $\mathcal{F}$ as $(\mathcal{F}(\rho))(x) = \gamma(x) = \Pi_\rho(\hat{\theta}_\rho^{-1}([0, x]))$. Lemma 11 will show that $\mathcal{F}$ maps $\mathcal{P}$ into itself.

Now to prove the above theorem we need to show that $\mathcal{F}$ has a fixed point, i.e $\mathcal{F}(\rho) = \rho$.

Theorem 6 (Schauder Fixed Point Theorem)**.** *Suppose $\mathcal{F}(\mathcal{P}) \subset \mathcal{P}$. If $\mathcal{F}(\cdot)$ is continuous, and $\mathcal{F}(\mathcal{P})$ is contained in a convex and compact subset of $\mathcal{P}$, then $\mathcal{F}(\cdot)$ has a fixed point.*

In next section, we show that the mapping $\mathcal{F}$ satisfies the conditions of the above theorem, and hence it has a fixed point. Note that $\mathcal{P}$ is a convex set. Therefore, we just need to show that the other two conditions are satisfied.

VI. MFE Existence: Proof

A. Continuity of the map $\mathcal{F}$

To prove the continuity of mapping $\mathcal{F}$, we first show that θ^* and Π^* are continuous mappings. To that end, we will show that for any sequence $\rho_n \to \rho$ in uniform norm, we have $\theta^*(\rho_n) \to \theta^*(\rho)$ in w-norm and $\Pi^*(\rho_n) \Rightarrow \Pi^*(\rho)$ (where $\Rightarrow$ denotes weak convergence). Finally, we use the continuity of θ^* and Π^* to prove that $\mathcal{F}(\rho_n) \to \mathcal{F}(\rho)$.

Step 1: Continuity of θ^*

Theorem 7. *The map θ^* is continuous.*

Proof: Define the map $V^* : \mathcal{P} \mapsto \mathcal{V}$ that takes ρ to $\hat{V}_\rho(\cdot)$. We begin by showing that $\|\hat{\theta}_{\rho_1} - \hat{\theta}_{\rho_2}\|_w \leq K\|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w$, which means that the continuity of the map V^* implies the continuity of the map θ^*.

Then we show two simple properties of the Bellman operator. The first is that for any $\rho \in \mathcal{P}$ and $f_1,\ f_2 \in \mathcal{V}$,

$$\|T_\rho f_1 - T_\rho f_2\|_w \leq \hat{K}\|f_1 - f_2\|_w \tag{14}$$

for some large $\hat{K}$, independent of ρ.

Second, let T_{ρ_1} and T_{ρ_2} be the Bellman operators corresponding to $\rho_1,\ \rho_2 \in \mathcal{P}$ and let $f \in \mathcal{V}$. We show that

$$\|T_{\rho_1} f - T_{\rho_2} f\|_w \leq 2(M - 1)K_1\|f\|_w\|\rho_1 - \rho_2\|. \tag{15}$$

We then have

$$\begin{aligned}
\|T^j_{\rho_1}\hat{V}_{\rho_2} - T^j_{\rho_2}\hat{V}_{\rho_2}\|_w &\leq \|T^j_{\rho_1}\hat{V}_{\rho_2} - T^{j-1}_{\rho_1}T_{\rho_2}\hat{V}_{\rho_2}\|_w \\
&\quad + \|T^{j-1}_{\rho_1}T_{\rho_2}\hat{V}_{\rho_2} - T^{j-2}_{\rho_1}T^2_{\rho_2}\hat{V}_{\rho_2}\|_w + \cdots \\
&\quad + \|T_{\rho_1}T^{j-1}_{\rho_2}\hat{V}_{\rho_2} - T^j_{\rho_2}\hat{V}_{\rho_2}\|_w \\
&\leq \hat{K}^{j-1}\|T_{\rho_1}\hat{V}_{\rho_2} - T_{\rho_2}\hat{V}_{\rho_2}\|_w + \cdots \\
&\quad + \|T_{\rho_1}T^{j-1}_{\rho_2}\hat{V}_{\rho_2} - T^j_{\rho_2}\hat{V}_{\rho_2}\|_w \qquad (16) \\
&\leq (\hat{K}^{j-1} + \cdots + 1)\|T_{\rho_1}\hat{V}_{\rho_2} - T_{\rho_2}\hat{V}_{\rho_2}\|_w \\
&\leq 2(M - 1)K\|\rho_1 - \rho_2\|(\hat{K}^{j-1} + \cdots + 1)\|\hat{V}_{\rho_2}\|_w \qquad (17)
\end{aligned}$$

Here, (16) and (17) follow from (14) and (15), respectively.

Now, let j be such that $T^j_{\rho_1}$ is an α-contraction. Note that Statement 1 of Lemma 2 implies that such a $j < \infty$ exists. Then we have

$$\begin{aligned}
&\|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w = \|T^j_{\rho_1}\hat{V}_{\rho_1} - T^j_{\rho_2}\hat{V}_{\rho_2}\|_w \\
&\leq \|T^j_{\rho_1}\hat{V}_{\rho_1} - T^j_{\rho_1}\hat{V}_{\rho_2}\|_w + \|T^j_{\rho_1}\hat{V}_{\rho_2} - T^j_{\rho_2}\hat{V}_{\rho_2}\|_w \\
&\implies (1 - \alpha)\|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w \leq \|T^j_{\rho_1}\hat{V}_{\rho_2} - T^j_{\rho_2}\hat{V}_{\rho_2}\|_w \qquad (18)
\end{aligned}$$

Finally, from (17) and (18), we get

$$\begin{aligned}
&\|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w \\
&\leq \frac{2(m-1)K(\hat{K}^{j-1} + \cdots + 1)\|\rho_1 - \rho_2\|}{1 - \alpha}\|\hat{V}_{\rho_2}\|_w \\
&\leq \frac{2(m-1)K(\hat{K}^{j-1} + \cdots + 1)\|\rho_1 - \rho_2\|}{1 - \alpha} \\
&\quad \times (\|\hat{V}_{\rho_1}\|_w + \|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w).
\end{aligned}$$

Therefore, if $\frac{2(m-1)K(\hat{K}^{j-1}+\cdots+1)}{1-\alpha}\|\rho_1 - \rho_2\| < \frac{1}{2}$, then

$$\begin{aligned}
&\|\hat{V}_{\rho_1} - \hat{V}_{\rho_2}\|_w \\
&\leq \frac{4(m-1)K(\hat{K}^{j-1} + \cdots + 1)}{1 - \alpha}\|\hat{V}_{\rho_1}\|_w\|\rho_1 - \rho_2\|
\end{aligned}$$

Hence, the maps V^* and θ^* are continuous. ∎

Step 2: Continuity of the map Π^*

Let $\Pi_{\rho,\theta}(.)$ be the invariant distribution generated by any θ. Recall that Π^* takes $\rho \in \mathcal{P}$ to probability measure $\Pi_\rho(.) = \Pi_{\rho,\hat{\theta}_\rho}(.)$. First, we show that $\Pi_{\rho,\theta}(.) \in \Omega$, where Ω is the space of absolutely continuous measures (with respect to Lebesgue measure) on $\mathbb{R}^+$.

Lemma 8. *For any $\rho \in \mathcal{P}$ and any $\theta \in \Theta$, $\Pi_{\rho,\theta}(\cdot)$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^+$.*

Proof: $\Pi_{\rho,\theta}(\cdot)$ can be expressed as the invariant queue-length distribution of the dynamics

$$q \to \begin{cases} Q' + A & \text{with probability } \beta \\ R & \text{with probability } (1 - \beta), \end{cases}$$

where $A \sim \Phi$ and $R \sim \Psi$, and Q' is a random variable with distribution generated by the conditional probabilities

$$\begin{aligned}
\mathbb{P}(Q' = q|q) &= 1 - p_\rho(\hat{\theta}(q)) \\
\mathbb{P}(Q' = (q-1)^+|q) &= p_\rho(\hat{\theta}(q))
\end{aligned}$$

Let Π' be the distribution of Q'. Then for any Borel set B, Π can be expressed using the convolution of Π' and Φ :

$$\Pi_{\rho,\theta}(B) = \beta \int_{-\infty}^{\infty} \Phi(B-y) d\Pi'(y) + (1-\beta)\Psi(B). \quad (19)$$

If B is a Lebesgue null-set, then so is $B - y \ \ \forall y$. So, $\Phi(B - y) = 0$ and $\Psi(B) = 0$ and therefore $\Pi_\rho(B) = 0$. ■

We now develop a useful characterization of $\Pi_{\rho,\theta}$. Let

$$\Upsilon^{(k)}_{\rho,\theta}(B|q) = \mathbb{P}(Q_k \in B | \text{no regeneration}, \ Q_0 = q)$$

be the distribution of queue length Q_k at time k induced by the transition probabilities (8) conditioned on the event that $Q_0 = q$ and that there are no regenerations until time k. We can now express the invariant distribution $\Pi_{\rho,\theta}(\cdot)$ in terms of $\Upsilon^{(k)}_{\rho,\theta}(\cdot|q)$ as in the following lemma.

Lemma 9. *For any bid distribution $\rho \in \mathcal{P}$ and for any stationary policy $\theta \in \Theta$, the Markov chain described by the transition probabilities in (8) has a unique invariant distribution $\Pi_{\rho,\theta}(\cdot)$ given by,*

$$\Pi_{\rho,\theta}(B) = \sum_{k \geq 0} (1-\beta)\beta^k \mathbb{E}_\Psi(\Upsilon^{(k)}_{\rho,\theta}(B|Q)), \quad (20)$$

where $\mathbb{E}_\Psi(\Upsilon^{(k)}_{\rho,\theta}(B|Q)) = \int \Upsilon^{(k)}_{\rho,\theta}(B|q) d\Psi(q)$.

Proof: $\Upsilon^{(k)}_{\rho,\theta}(B|q)$ is the queue length distribution assuming no regeneration has happened yet, and the regeneration event occurs with probability β independently of the rest of the system. It is then easy to find $\Pi_{\rho,\theta}(B)$ in terms of $\Upsilon^{(k)}_{\rho,\theta}(B|q)$ by simply using the properties of the conditional expectation, and the theorem follows. Note that in $\mathbb{E}_\Psi(\Upsilon^{(k)}_{\rho,\theta}(B|Q))$, the random variable is the initial condition of the queue, as generated by Ψ. ■

We shall now prove the continuity of Π^* in ρ.

Theorem 10. *The mapping $\Pi^* : \mathcal{P} \mapsto \Omega$ is continuous.*

Proof: By Portmanteau theorem [18], we only need to show that for any sequence $\rho_n \to \rho$ in w-norm and any open set B, $\liminf_{n\to\infty} \Pi_{\rho_n}(B) \geq \Pi_\rho(B)$. By Fatou's lemma,

$$\begin{aligned}
&\liminf_{n\to\infty} \Pi_{\rho_n}(B) \\
&= \liminf_{n\to\infty} \sum_{k=0}^{\infty} (1-\beta)\beta^k \mathbb{E}_{\Psi_R}[\Upsilon^{(k)}_{\rho_n}(B|Q)] \\
&\geq \sum_{k=0}^{\infty} (1-\beta)\beta^k \mathbb{E}_{\Psi_R}[\liminf_{n\to\infty} \Upsilon^{(k)}_{\rho_n}(B|Q)] \quad (21)
\end{aligned}$$

where $Q \sim \Psi_R$. Let $\Upsilon^{(k)}_\rho = \Upsilon^{(k)}_{\rho,\hat{\theta}_\rho}$. We prove by induction that $\liminf_{n\to\infty} \Upsilon^{(k)}_{\rho_n}(B|q) \geq \Upsilon^{(k)}_\rho(B|q)$ for every $q \in \mathbb{R}^+$, and the proof follows. ■

Step 3: Continuity of the mapping $\mathcal{F}$

Now, using the results from Step 1 and Step 2, we establish continuity of the mapping $\mathcal{F}$. First, we show that $\mathcal{F}(\rho) \in \mathcal{P}$.

Lemma 11. *For any $\rho \in \mathcal{P}$, let $\gamma(x) = (\mathcal{F}(\rho))(x) = \Pi_\rho(\hat{\theta}_\rho^{-1}([0,x])), x \in \mathbb{R}^+$. Then, $\gamma \in \mathcal{P}$.*

Proof: From the definition of Π_ρ, it is easy to see that γ is a distribution function. Since $\hat{\theta}_\rho$ is continuous and strictly increasing function as shown in Lemma 4, $\hat{\theta}_\rho^{-1}(\{x\})$ is either empty or a singleton. Then, from Lemma 8, we get that $\Pi_\rho(\hat{\theta}_\rho^{-1}(\{x\})) = 0$. Together, we get that $\gamma(x)$ has no jumps at any x and hence it is continuous.

To complete the proof, we need to show that the expected bid under $\gamma(.)$ is finite. In order to do this, we construct a new random process $\tilde{Q}_k$ that is identical to the original queue length dynamics Q_k, except that it never receives any service. We show that this process stochastically dominates the original, and use this property to bound the mean of the original process by a finite quantity independent of ρ. ■

We now have the main theorem.

Theorem 12. *The mapping $\mathcal{F} : \mathcal{P} \mapsto \mathcal{P}$ given by $(\mathcal{F}(\rho))(x) = \Pi_\rho(\hat{\theta}_\rho^{-1}([0,x]))$ is continuous.*

Proof: Let $\rho_n \to \rho$ in uniform norm. From previous steps, we have $\hat{\theta}_{\rho_n} \to \hat{\theta}_\rho$ in w-norm and $\Pi_{\rho_n} \Rightarrow \Pi_\rho$. Then, using Theorem 5.5 of Billingsley [18], one can show that the push-forwards also converge:

$$\Pi_{\rho_n}(\hat{\theta}_{\rho_n}^{-1}(\cdot)) \Rightarrow \Pi_\rho(\hat{\theta}_\rho^{-1}(\cdot)).$$

Then, $\mathcal{F}(\rho_n)$ converges point-wise to $\mathcal{F}(\rho)$ as it is continuous at every x, i.e., $(\mathcal{F}(\rho_n))(x) \to (\mathcal{F}(\rho_n))(x)$ for all $x \in \mathbb{R}^+$.

Using standard arguments it can be shown that in the norm space $\mathcal{P}$, point-wise convergence implies convergence in uniform norm. This completes the proof. ■

B. $\mathcal{F}(\mathcal{P})$ contained in a compact subset of $\mathcal{P}$

We show that the closure of the image of the mapping $\mathcal{F}$, denoted by $\overline{\mathcal{F}(\mathcal{P})}$, is compact in $\mathcal{P}$. As $\mathcal{P}$ is a normed space, sequential compactness of any subset of $\mathcal{P}$ implies that the subset is compact. Hence, we just need to show that $\overline{\mathcal{F}(\mathcal{P})}$ is sequentially compact. Sequential compactness of a set $\overline{\mathcal{F}(\mathcal{P})}$ means the following: if $\{\rho_n\} \in \overline{\mathcal{F}(\mathcal{P})}$ is a sequence, then there exists a subsequence $\{\rho_{n_j}\}$ and $\rho \in \overline{\mathcal{F}(\mathcal{P})}$ such that $\rho_{n_j} \to \rho$. We use Arzelà-Ascoli theorem and uniform tightness of the measures in $\mathcal{F}(\mathcal{P})$ to show the sequential compactness. The version that we will use is stated below:

Theorem 13 (Arzelà-Ascoli Theorem)**.** *Let X be a σ-compact metric space. Let $\mathcal{G}$ be a family of continuous real valued functions on X. Then the following two statements are equivalent:*

1) *For every sequence $\{g_n\} \subset \mathcal{G}$ there exists a subsequence g_{n_j} which converges uniformly on every compact subset of X.*
2) *The family $\mathcal{G}$ is equicontinuous on every compact subset of X and for any $x \in X$, there is a constant C_x such that $|g(x)| < C_x$ for all $g \in \mathcal{G}$.*

Suppose a family of functions $\mathcal{D} \subseteq \mathcal{P}$ satisfies the equivalent conditions of the Arzelá-Ascoli theorem and in addition satisfy the uniform tightness property, i.e., $\forall \epsilon > 0$ there exists and x_ϵ such that for all $f \in \mathcal{D}$ $1 \geq f(x_e) \geq 1 - \epsilon$. Then, for any sequence $\{\rho_n\} \subset \mathcal{D}$, there exists a subsequnce $\{\rho_{n_j}\}$ that converges uniformly on every compact set to a continuous increasing function ρ on $\mathbb{R}^+$. As $\mathcal{D}$ is uniformly tight it can be shown that ρ_{n_j} converges uniformly to ρ and that $\rho \in \mathcal{P}$. Therefore, $\overline{\mathcal{D}}$ is sequentially compact in the topology of uniform norm.

In the following, we show that $\mathcal{F}(\mathcal{P})$ satisfies uniform tightness property and condition 2 in Arzelá-Ascoli theorem. First verifying the conditions of Arzelá-Ascoli theorem, note that the functions in consideration are uniformly bounded by 1. To prove equicontinuity, consider an $\gamma = \mathcal{F}(\rho)$ and let $x > y$.

$$\begin{aligned}\gamma(x) - \gamma(y) &= \Pi_\rho(\theta_\rho(q) \le x) - \Pi_\rho(\theta_\rho(q) \le y) \\ &= \Pi_\rho(y < \theta_\rho(q) \le x) \qquad (22)\end{aligned}$$

Lemma 14. *For any interval $[a, b]$, $\Pi_\rho([a, b]) < c \cdot (b - a)$, for some large enough c.*

Proof: The proof follows easily from our characterization of Π_ρ in terms of $\Upsilon_\rho^{(k)}$. ■

The above lemma and equation (22) imply that $\gamma(x) - \gamma(y) \le c(\theta_\rho^{-1}(x) - \theta_\rho^{-1}(y))$. To show equicontinuity, it is enough to show that $\limsup_{y\uparrow x} \frac{\gamma(x)-\gamma(y)}{x-y} \le K(x)$ for some K independent of ρ. This property follows from our characterization of the optimal bid function.

Finally, we note that $\mathcal{P}$, and hence $\mathcal{F}(\mathcal{P})$, is uniformly tight. This is due to the fact that the expectation of the bid distributions is bounded uniformly. An application of Markov inequality will give uniform tightness.

VII. Approximation Results: PBE and MFE

In this section we prove that the mean field policy is an ϵ-Nash equilibrium. We have the following theorem:

Theorem 15. *Let $(\rho, \hat{\theta}_\rho)$ constitute an MFE. Suppose at time 0 the queue length of the users is set independently across users according to Π_ρ; and that their initial belief is also consistent. Also, suppose that all queues except queue 1 play the MFE policy $\hat{\theta}_\rho$. Then, for any policy θ^N of queue 1 that may be history dependent and any $q \in \mathbb{R}^+$, we have*

$$\limsup_{N\to\infty} V^N_{1,\mu_{1,0}}(q; \hat{\theta}_\rho, (\hat{\theta}_\rho)_{-1}) - V^N_{1,\mu_{1,0}}(q; \theta^N, (\hat{\theta}_\rho)_{-1}) \le 0,$$

where $\mu_{1,0} = \Pi_\rho$ and the superscript N has been added to explicitly indicate the dependence on the number of cells.

The main idea behind the proof is a result called *propagation of chaos,* and it identifies conditions under which any finite subset of the state variables are independent from each other. We state this result now in our context.

Lemma 16 (Propagation of chaos)**.** *For any fixed indices $i_1, \dots, i_k$, let $\mathcal{L}(Q^n_{i_1}(t), \dots, Q^n_{i_k}(t)$ denote the probability law of the k-tuple of corresponding queues in the MN-queue system, at time t. Suppose that $\mathcal{L}((Q^N_{i_1}(0), \dots, Q^N_{i_k,0}(0)) = \otimes^k \Pi_\rho$, where (ρ, θ_ρ) is the solution to the MFE equation. Also, suppose that all queues are following mean field equilibrium strategy. Then for any $T > 0$, we have*

$$\mathcal{L}(Q^N_{i_1}(T), \dots, Q^N_{i_k}(T) \Rightarrow \otimes^k \Pi_\rho,$$

as $N \to \infty$.

Proof: We shall only consider the case $k = 2$; the proof of the general case is similar. We follow the proof of Theorems 4.1 and 5.1 in Graham and Meleard [13]. The proof is divided into two parts; the first part proves that for any two agents i and j,

$$\|\mathcal{L}(Q^N_i, Q^N_j) - \mathcal{L}(Q^N_i) \otimes \mathcal{L}(Q^N_j)\|_D \to 0,$$

where the subscript D refers to the total variation norm. In the second, we show that $\mathcal{L}(Q^N_i) \Rightarrow \Pi_\rho$. Both parts rely on studying interaction graphs, defined in [13], which characterize the amount of interactions that any finite subset of agents may have had in the past. ■

The proof of Theorem 15 is as follows. Suppose we start at time $t = 0$ with queue length of agent 1 being $Q_1(0)$. We can choose a time T large enough so that the value added by auctions occurring after time T is less than ϵ, due to discounting. Thus, the difference in value contributed by these auctions, when using policy θ^N and $\hat{\theta}_\rho$ can be bounded by 2ϵ, and we can restrict attention to the first T auctions.

Using ideas similar to Lemma 16, we show that probability of the event E^N that other agents that interact with agent 1 at time t have never been influenced by agent 1 goes to 1 as N becomes large. Thus, the belief of distribution of queue lengths of other agents encountered converges to Π_ρ according to Lemma 16. Then we can show that for any $\epsilon > 0$ and N,

$$V^N_{1,\mu_{1,0}}(q; \theta_\rho, (\theta_\rho)_{-1}) - V^N_{1,\mu_{1,0}}(q; \theta^N, (\theta_\rho)_{-1}) \le \epsilon,$$

which yields the desired result.

VIII. Simulation Results

We now turn to computing the MFE distribution. For simplicity of simulation, we truncate and discretize both state and bid spaces. The new state space is $\mathcal{S} = \{0.01m, 0 \le m \le 2000\}$, while the bid space is $\mathcal{X} = \{0.15m, 0 \le m \le 3000\}$. The job arrival and regeneration distributions are both chosen to be uniform over interval $[0, 1]$. The service rate of each base station is assumed to be 5 units per time slot. Finally, the holding cost function is chosen as $C(q) = q^2$.

We develop an algorithm for our setting inspired by the computational approach in [11]. Let $\rho_0(x) = \min\{0.001x, 1\}, x \in \mathcal{X}$ and $\Pi_0 = \Psi$. For every positive integer n, do the following:

1) Compute the optimal value function, $\hat{V}_n$, which is the unique solution to the following equation,

$$\hat{V}_n(q) = C(q) + \beta\mathbb{E}_A[\hat{V}_n(q+A)] - \sum_{x \le \beta\Delta\hat{V}_n(q), x\in\mathcal{X}} p_{\rho_n}(x),$$

where $\Delta\hat{V}_n(q) = \mathbb{E}_A[\hat{V}_n(q+A)] - \mathbb{E}_A[\hat{V}_n((q-1)^+ + A)]$. We apply value iteration method (Section 6.10 [17]) to compute an approximate solution to the above equation.

2) Compute the optimal bid function, $\hat{\theta}_n$ as

$$\hat{\theta}_n(q) = \beta\mathbb{E}_A[\hat{V}_n(q + A) - \hat{V}_n((q - 1)^+ + A)].$$

3) Compute steady state distribution, Π_n as follows: Let $p(q) = p_{\rho_n}(\hat{\theta}_n(q))$. Then,

$$\Pi_n[q] = (1 - \beta)\Psi(q) + \beta \sum_{q'\in\mathcal{S}} \Pi_{n-1}(q') \times \left(p(q')\mathbb{E}_A[\mathbf{1}_{((q'-5)^+ + A = q)}] + (1 - p(q'))\mathbb{E}_A[\mathbf{1}_{(q'+A=q)}]\right).$$

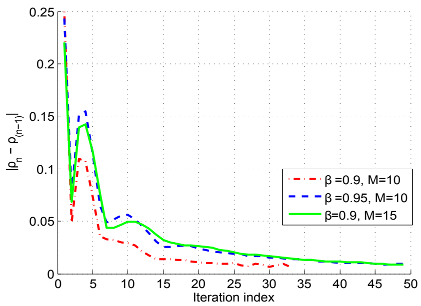

Fig. 2. Convergence to MFE bid distribution

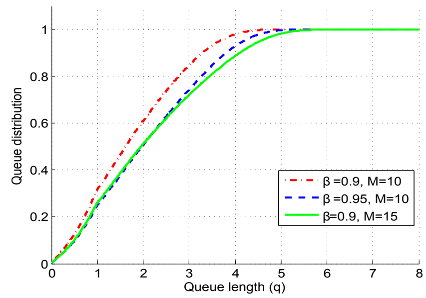

Fig. 3. MFE queue length distribution

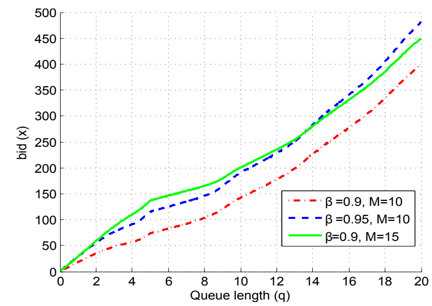

Fig. 4. MFE optimal bid function

4) Finally, compute bid distribution, $\rho_{n+1}(x) = \Pi_n[\hat{\theta}_n^{-1}([0,x])]$. If $||\rho_n - \rho_{n+1}|| < \epsilon$, then $\rho = \rho_{n+1}$ and exit. Otherwise, set $n = n+1$ and go to Step 1.

If the algorithm converges, then its output distribution, ρ, is an approximation of the MFE bid distribution.

We simulated the algorithm for three set of parameters: 1. ($\beta = 0.9, M = 10$), 2. ($\beta = 0.95, M = 10$) and 3. ($\beta = 0.9, M = 15$). Also, we chose the accuracy parameter $\epsilon = 0.008$. Figure 2 shows that the algorithm converges in less than 50 iterations in all three cases. In each iteration, Step 1 (value iteration) is the most computationally intensive. It converges in 80 recursions, with each recursion having to update $|\mathcal{S}|$ number of variables, and with each variable update requiring at most $|\mathcal{X}|$ number of arithmetic operations. All together, the computational complexity of each iteration is in the order of $80 \times |\mathcal{S}| \times |\mathcal{X}|$ arithmetic operations.

The queue length distributions at MFE are shown in Figure 3. We observe that the distribution curves exhibit a rightward shift with increase in β or M. Note that larger β makes queues live longer without regeneration, while higher M reduces each individual's average service rate. Hence, the queues get longer on average. We show the optimal bid functions at MFE in Figure 4. As expected from our analysis, the bid functions are monotonically increasing in queue length.

Discussion on Implementation

The ease of computation of the MFE immediately suggests a simple implementation scheme. Suppose that the base stations were to calculate the empirical bid distribution at each time instant, and return it to the users. The users play their best response to this bid distribution. Complexity of computing the best response is low, since value iteration by each user has to be done only for a singe queue state. The base stations combine all the bids to create a new empirical bid distribution. Such a proceeding is essentially identical to the algorithm that we employed above, and would converge in a similar fashion.

IX. Conclusion

In this paper, we have explored the question of whether it is possible to design an incentive compatible scheduling policy that has the attractive properties of an LQF service regime. We used a mean field framework to show that as the number of agents in the system becomes large, this objective can indeed be fulfilled using a second price auction at each server. In reality, such auctions would have to be conducted periodically (say every five minutes), and the servers would actually be a set of OFDM channel at each base station to be allocated to the highest bidders. There would also be different classes of applications having different cost functions and arrival rates. It is straightforward to incorporate these extensions to our model, and we plan on conducting a trial in our Smart Phone Laboratory in the future.

References

[1] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," in *IEEE Transactions on Automatic Control*, vol. 37, 1992, pp. 1936–1948.

[2] X. Lin and N. B. Shroff, "Joint rate control and scheduling in multihop wireless networks," in *in Proceedings of IEEE Conference on Decision and Control*, 2004, pp. 1484–1489.

[3] A. Eryilmaz and R. Srikant, "Joint congestion control, routing and mac for stability and fairness in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 1514–1524, 2006.

[4] M. J. Neely, E. Modiano, and C. ping Li, "Fairness and optimal stochastic control for heterogeneous networks," in *Proceedings of IEEE INFOCOM*, 2005, pp. 1723–1734.

[5] S. Ha, S. Sen, C. Joe-Wong, Y. Im, and M. Chiang, "Tube: time-dependent pricing for mobile data," in *SIGCOMM*, 2012, pp. 247–258.

[6] V. Krishna, *Auction theory*. Academic press, 2009.

[7] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japan Journal of Mathematics*, 2007.

[8] H. Tembine, J.-Y. Le Boudec, R. El-Azouzi, and E. Altman, "Mean field asymptotics of Markov decision evolutionary games and teams," in *GameNets*, 2009, pp. 140–150.

[9] S. Adlakha, R. Johari, and G. Weintraub, "Equilibria of dynamic games with many players: Existence, approximation, and market structure," 2010, working paper, arXiv:1011.5537 [cs.GT].

[10] V. Borkar and R. Sundaresan, "Asymptotics of the invariant measure in mean field models with jumps," in *49th Annual Allerton Conference on Communication, Control, and Computing*, 2011, pp. 1258–1263.

[11] K. Iyer, R. Johari, and M. Sundararajan, "Mean field equilibria of dynamic auctions with learning," *ACM SIGecom Exchanges*, vol. 10, no. 3, pp. 10–14, 2011.

[12] J. Xu and B. Hajek, "The supermarket game," in *ISIT*, 2012.

[13] 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.

[14] R. E. Strauch, "Negative dynamic programming," *The Annals of Mathematical Statistics*, vol. 37, no. 4, pp. 871–890, 1966.

[15] S. P. Meyn, R. L. Tweedie, and P. W. Glynn, *Markov chains and stochastic stability*. Cambridge University Press, 2009, vol. 2.

[16] M. Manjrekar, V. Ramaswamy, and S. Shakkottai, "A mean field game approach to scheduling in cellular systems," Tech. Report available at http://www.ece.tamu.edu/~sshakkot/mean-field-tech.pdf.

[17] M. L. Puterman, *Markov decision processes: Discrete stochastic dynamic programming*. John Wiley & Sons, Inc., 1994.

[18] P. Billingsley, *Convergence of probability measures*. Wiley-Interscience, 2009, vol. 493.