Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Oblivious Equilibrium for Large-Scale Stochastic Games with Unbounded Costs

Sachin Adlakha, Ramesh Johari, Gabriel Weintraub, and Andrea Goldsmith

Abstract— **We study stochastic dynamic games with a large number of players, where players are coupled via their cost functions. A standard solution concept for stochastic games is *Markov perfect equilibrium* (MPE). In MPE, each player's strategy is a function of its own state as well as the state of the other players. This makes MPE computationally prohibitive as the number of players becomes large. An approximate solution concept called *oblivious equilibrium* (OE) was introduced in [1], where each player's decision depends only on its own state and the "long-run average" state of other players. This makes OE computationally more tractable than MPE. It was shown in [1] that, under a set of assumptions, as the number of players become large, OE closely approximates MPE. In this paper we relax those assumptions and generalize that result to cases where the cost functions are unbounded. Furthermore, we show that under these relaxed set of assumptions, the OE approximation result can be applied to large population linear quadratic Gaussian (LQG) games [2].**

I. INTRODUCTION

In this paper, we study stochastic games with a large number of players. Such games are used to model complex dynamical systems such as wireless networks [3], [4], industry dynamics with many firms [5], etc. In such games, the players typically have competing objectives. A common equilibrium notion for such games is *Markov perfect equilibrium*. In MPE, each player minimizes its individual cost by choosing a strategy that is a function of the current state of all the players.

MPE suffers from at least two drawbacks. First, as an equilibrium concept, it is implausible in settings where many agents interact with each other; MPE requires the agents to be aware of the state evolution of all other agents. Second, MPE can be computationally intractable. MPE is typically obtained numerically using dynamic programming (DP) algorithms [7]. Thus, as the number of players increases, the cost of computing the strategy and maintaining the state of all the players becomes prohibitively large [8]. Several techniques have been proposed in the literature to deal with the complexity of large scale systems [9], [10], [15].

Recently, a scheme for approximating MPE for such large scale games was proposed in [1], via a solution concept called *oblivious equilibrium*. In oblivious equilibrium, a player optimizes given only the long-run *average* statistics of other players, rather than the entire instantaneous vector of its competitors' state. OE resolves both of the difficulties raised above: in OE, a player is reacting to far simpler aggregate statistics of the behavior of other players. Further, OE computation is significantly simpler than MPE computation, since each player only needs to solve a one-dimensional dynamic program.

Under what conditions will OE approximate MPE? When there are a large number of players in a game, an individual player can optimize based on the average behavior of it competitors provided (a) the actual behavior of the competitors is close to the average behavior, and (b) any deviation from the average behavior has "small" impact on the cost of the individual player. It is intuitive to believe that as the number of players in a game becomes large, the actual behavior would approach its mean by a law of large numbers effect. To measure the impact of any deviation from the average behavior, the authors in [1] defined a "light-tail" condition. Informally, this condition implies that the effect of a small perturbation in the instantaneous state of the competitors has a small effect on the cost of a player. It is reasonable to expect that under such a condition, if players make decisions based only on the long-run average, they should achieve near-optimal performance. Indeed, it is established in [1] that under a reasonable set of technical conditions (including the "light-tail" condition), OE is a good approximation to MPE for industry dynamic models with many firms; formally, this is called the *asymptotic Markov equilibrium* (AME) property.

As presented in [1], the main approximation result is tailored to the class of firm competition models presented there. In [6], using the methods of [1], the authors isolated a set of parsimonious assumptions for a general class of stochastic games, under which OE is a good approximation to MPE. They also study the case of non-uniform players with heterogeneous cost functions.

However, the results in [1], [6] are based on the assumption that the cost functions are uniformly bounded over states and actions. This is a restrictive assumption; for example, typical cost functions used in decentralized control are unbounded—e.g., quadratic cost. In this paper, we remove this boundedness assumption on the cost functions; this change necessitates an alternate approach to establish that OE is a good approximation to MPE. We show a general result under this assumption, then apply our result to a class of games including games with quadratic cost and linear dynamics. The latter result is a generalization of a similar result derived by [2].

The rest of the paper is organized as follows. In section II, we outline our model of stochastic game, describe our

S. Adlakha and A. Goldsmith are with the Department of Electrical Engineering, Stanford University. adlakha@stanford.edu, andrea@wsl.stanford.edu

R. Johari is with the Department of Management Science and Engineering, Stanford University. ramesh.johari@stanford.edu

G. Weintraub is with the Columbia Business School, Columbia University. gweintraub@columbia.edu

978-1-4244-3124-3/08/$25.00 ©2008 IEEE

notation and define the oblivious equilibrium. In section III, we introduce the asymptotic Markov equilibrium (AME) property, formally define the light-tail condition and introduce our assumptions on the cost function. Section IV details the proof of the main theorem. In section V, we show that linear-quadratic games can be reduced to a special case of our model. Section VI concludes the paper.

II. MODEL, DEFINITIONS AND NOTATION

We consider an m-player stochastic game evolving over discrete time periods with an infinite horizon. The discrete time periods are indexed by non-negative integers $t \in \mathbb{N}$. The state of a player i at time t is denoted by $x_{i,t} \in \mathcal{X}$, where $\mathcal{X}$ is a discrete subset of $\mathbb{R}$. We assume that the state evolution of a player i depends only on its own current state and the action it takes. This can be represented by a conditional probability mass function (pmf)

$$x_{i,t+1} \sim h\left(x \mid x_{i,t}, a_{i,t}\right), \quad (1)$$

where $a_{i,t}$ is the action taken by the player i at time t. We denote the set of actions available to a player by $\mathcal{A}$; we assume this is a discrete subset of Euclidean space $\mathbb{R}$.

The single period cost to a player i is given as $c\left(x_{i,t}, a_{i,t}, \boldsymbol{x}_{-i,t}\right)$. Here $\boldsymbol{x}_{-i,t}$ is the state of all players except player i at time t. Note that the cost to player i does not depend on the actions taken by other players. Furthermore, we assume that the cost function is independent of the identity of other players. That is, it only depends on the current state $x_{i,t}$ of player i, the total number m of players at any time, and the fraction $f_{-i,t}^{(m)}(y)$, which is the fraction of players excluding player i, that have their state as y. In other words, we can write the cost function as $c\left(x_{i,t}, a_{i,t}, f_{-i,t}^{(m)}, m\right)$, where $f_{-i,t}^{(m)}$ can be expressed as

$$f_{-i,t}^{(m)}(y) \triangleq \frac{1}{m-1} \sum_{j \neq i} \mathbf{1}_{\{x_{j,t}=y\}}. \quad (2)$$

From equation (1) and the definition of the cost function, we note that the players are coupled via their cost functions only.

Each player i chooses an action $a_{i,t} = \mu_i^{(m)}(x_{i,t}, f_{-i,t}^{(m)})$ to minimize its expected present value. Note that the policy $\mu_i^{(m)}$ depends on the total number of players m because of the underlying dependence of the cost function on m. Let $\boldsymbol{\mu}^{(m)}$ be the vector of policies of all players, and $\boldsymbol{\mu}_{-i}^{(m)}$ be the vector of policies of all players except player i. We define $V(x, f, m|\mu_i^{(m)}, \boldsymbol{\mu}_{-i}^{(m)})$ to be the expected net present value for player i with current state x, if the current aggregate state of players other than i is f, given that i follows the policy $\mu_i^{(m)}$ and the policy vector of players other than i is given by $\boldsymbol{\mu}_{-i}^{(m)}$. In particular, we have

$$\begin{aligned} V(x, f, m \mid \mu_i^{(m)}, \boldsymbol{\mu}_{-i}^{(m)}) &\triangleq \\ \mathbb{E}\Bigg[\sum_{t=0}^{\infty} \beta^t c(x_{i,t}, a_{i,t}, f_{-i,t}^{(m)}, m) &\,\Bigg|\, x_{i,0} = x, f_{-i,0}^{(m)} = f; \mu_i^{(m)}, \boldsymbol{\mu}_{-i}^{(m)}\Bigg], \end{aligned} \quad (3)$$

where $0 < \beta < 1$ is the discount factor. Note that the random variables $(x_{i,t}, f_{i,t}^{(m)})$ depend on the policy vector $\boldsymbol{\mu}^{(m)}$ and the state evolution function h.

We focus on *symmetric Markov perfect equilibrium*, where all players use the same policy $\mu^{(m)}$. We thus drop the subscript i in the policy of a player i. Let $\mathcal{M}$ be the set of all policies available to a player. Note that this set also depends on the total number of players m.

Definition 1 (Markov Perfect Equilibrium): The vector of policies $\boldsymbol{\mu}^{(m)}$ is a *Markov perfect equilibrium* if for all i, x, and f we have

$$\inf_{\mu' \in \mathcal{M}} V\left(x, f, m \mid \mu', \boldsymbol{\mu}_{-i}^{(m)}\right) = V\left(x, f, m \mid \mu^{(m)}, \boldsymbol{\mu}_{-i}^{(m)}\right).$$

As the number of players becomes large, the MPE becomes computationally intractable. This is because the set of all policies grows exponentially in the number of players. However, if the coupling between the players is weak, it is possible that the players can choose their optimal action based solely on their own state and the average state of the other players. We expect that as the number of players becomes large, the changes in the players' states average out such that the state vector $f_{-i,t}^{(m)}$ is well approximated by its long run average. Thus, each player can find its optimal policy based solely on its own state and the long-run average aggregate state of the other players.

We therefore restrict attention to policies that are only a function of the player's own state, and an underlying constant aggregate distribution of the competitors. Such strategies are referred to as *oblivious strategies* since they do not take into account the complete state of the competitors at any time. Let us denote $\tilde{\mu}^{(m)}$ as an oblivious policy of a player i; we let $\tilde{\mathcal{M}}$ denote the set of all oblivious policies available to a player. This set also depends on the number of players m. Note that if all players use oblivious strategies, their states evolve as independent Markov chains. We make the following assumption regarding the Markov chain of each player playing an oblivious policy.

Assumption 1: The Markov chain associated with the state evolution of each player i playing an oblivious policy $\tilde{\mu}^{(m)}$ is positive recurrent, and reaches a stationary distribution $q^{(m)}$.

The stationary distribution depends on the number of players m because the oblivious policy depends on m. Let $\tilde{\boldsymbol{\mu}}^{(m)}$ be the vector of oblivious policies for all players, $\tilde{\mu}_i^{(m)}$ be the oblivious policy for a player i, and $\tilde{\boldsymbol{\mu}}_{-i}^{(m)}$ be the vector of oblivious policies of all player except the player i. For simplification of analysis, we assume that the initial state of a player i is sampled from the stationary distribution $q^{(m)}$ of its state Markov chain; without this assumption, the OE approximation holds only after sufficient mixing of the individual players' state evolution Markov chains. Given $\tilde{\boldsymbol{\mu}}_{-i}^{(m)}$, for a particular player i, the long-run average aggregate state of its competitors is denoted by $\tilde{f}_{-i}^{(m)}$, and is defined as

$$\tilde{f}_{-i}^{(m)}(y) \triangleq \mathbb{E}\left(f_{-i,t}^{(m)}(y)\right) = q^{(m)}(y) \tag{4}$$

Note that, $\tilde{f}_{-i}^{(m)}$ is completely determined by the state evolution function h and the oblivious policy $\tilde{\boldsymbol{\mu}}_{-i}^{(m)}$.

As with the case of symmetric MPE defined above, we assume that players use the same oblivious policy denoted by $\tilde{\mu}^{(m)}$. We thus drop the subscript i in the oblivious policy of a player i. We define the *oblivious value function* $\tilde{V}(x, m|\tilde{\mu}^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)})$ to be the expected net present value for a player i with current state x, if player i follows the oblivious policy $\tilde{\mu}^{(m)}$, and players other than i follow the oblivious policy vector $\tilde{\boldsymbol{\mu}}_{-i}^{(m)}$. Specifically, we have

$$\tilde{V}(x, m \mid \tilde{\mu}^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)}) \triangleq \\ \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t c\left(x_{i,t}, a_{i,t}, \tilde{f}_{-i}^{(m)}, m\right) \;\middle|\; x_{i,0} = x; \tilde{\mu}^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)}\right]. \tag{5}$$

Note that the expectation does not depend explicitly on the policies used by players other than i; this dependence only enters through the long-run average aggregate state $\tilde{f}_{-i}^{(m)}$. In particular, the state evolution is *only* due to the policy of player i. Using the oblivious value function, we define oblivious equilibrium as follows.

Definition 2 (Oblivious Equilibrium): The vector of policies $\tilde{\mu}$ represents an *oblivious equilibrium* if for all i, we have

$$\inf_{\mu' \in \tilde{\mathcal{M}}} \tilde{V}\left(x, m \mid \mu', \tilde{\boldsymbol{\mu}}_{-i}^{(m)}\right) = \tilde{V}\left(x, m \mid \tilde{\mu}^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)}\right), \; \forall x.$$

In this paper, we do not show the existence of Markov perfect equilibrium or of oblivious equilibrium. We assume that both the equilibrium points exist for the stochastic game under consideration [14].

Assumption 2: Markov perfect equilibrium and oblivious equilibrium exist for the stochastic game under consideration.

III. ASYMPTOTIC MARKOV EQUILIBRIUM AND THE LIGHT TAIL

As mentioned before, we would like to approximate MPE using OE. To formalize the notion under which OE approximates MPE, we define the asymptotic Markov equilibrium (AME) property. Intuitively, this property says that an oblivious policy is approximately optimal even when compared against Markov policies. Formally, the AME ensures that as number of players in the game becomes large, the approximation error between the expected net present value obtained by deviating from the oblivious policy $\tilde{\mu}^{(m)}$ and instead following the optimal (non-oblivious) policy goes to zero for each state x of the player.

Definition 3 (Asymptotic Markov Equilibrium): We say that a sequence of oblivious policies $\tilde{\boldsymbol{\mu}}_{-i}^{(m)}$ possesses the asymptotic Markov equilibrium (AME) property if for all x and i, we have

$$\lim_{m\to\infty} \mathbb{E}\left[V\left(x, f, m \mid \tilde{\mu}^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)}\right) - \inf_{\mu' \in \mathcal{M}} V\left(x, f, m \mid \mu', \tilde{\boldsymbol{\mu}}_{-i}^{(m)}\right)\right] = 0.$$

Notice that the expectation here is over f, which denotes the aggregate state of all players other than i. MPE requires the error to be zero for all (x, f), rather than in expectation; of course, in general, it will not be possible to find a single oblivious policy that satisfies the AME property for any f. In particular, in OE, actions taken by a player will perform poorly if the other players' state is far from the long-run average aggregate state. Thus, AME implies that the OE policy performs nearly as well as the non-oblivious best policy for those aggregate states of other players that occur with high probability.

In order to establish the AME property, we make some assumptions on the cost functions. For notational convenience, we drop the subscripts i, t whenever it does not lead to any ambiguity. Note that in oblivious equilibrium, we replace the actual distribution of the opponents' state $f_{-i,t}^{(m)}$ by its mean distribution $\tilde{f}_{-i}^{(m)}$. In order to measure the *distance* between the actual distribution of the state and its mean, we define a notion of norm on a distribution.

Definition 4 (1-g Norm): Given a function $g : \mathcal{X} \to [0, \infty)$, we define the $1-g$ norm of the distribution f as

$$\|f\|_{1-g} = \sum_y |f(y)| g(y) \tag{6}$$

The $1-g$ norm is a weighted norm where g is the weight function. Note that this function depends on the actual form of the cost function. For a given function g, the distance between the actual distribution $f_{-i,t}^{(m)}$ and its mean $\tilde{f}_{-i}^{(m)}$ is given by

$$\left\| f_{-i,t}^{(m)} - \tilde{f}_{-i}^{(m)} \right\|_{1-g} = \sum_y \left| f_{-i,t}^{(m)}(y) - \tilde{f}_{-i}^{(m)}(y) \right| g(y).$$

We now formally define the *light tail* condition.

Assumption 3 (Light Tail): Given any $\epsilon > 0$, there exists a state value $z > 0$, such that

$$\mathbb{E}\left[g(\tilde{U}^{(m)}) \mathbf{1}_{|\tilde{U}^{(m)}| > z} \mid \tilde{U}^{(m)} \sim \tilde{f}_{-i}^{(m)}\right] \leq \epsilon, \quad \forall m, \tag{7}$$

$$\mathbb{E}\left[g^2(\tilde{U}^{(m)}) \mathbf{1}_{|\tilde{U}^{(m)}| > z} \mid \tilde{U}^{(m)} \sim \tilde{f}_{-i}^{(m)}\right] \leq \epsilon, \quad \forall m. \tag{8}$$

Here $\tilde{U}^{(m)}$ is a random variable distributed according to $\tilde{f}_{-i}^{(m)}$. As mentioned before, $g(y)$ is a weight function for the state y. Thus, the light tail assumption requires that the weighted tail probability of the competitors goes to zero uniformly over m. Also, note that only the second condition needs to be checked; a straightforward application of Jensen's inequality then shows the first condition must hold. However, we retain both inequalities for clarity. Furthermore, if $\sup_y g(y) < \infty$, then the light tail condition is just an assumption on the tail of the mean distribution $\tilde{f}_{-i}^{(m)}$.

In order that the AME property holds, we would like that the cost functions are close to each other when the actual distribution $f_{-i,t}^{(m)}$ is close to its mean distribution $\tilde{f}_{-i}^{(m)}$. For this to hold, we impose a uniformity condition on the growth of the cost function. Specifically, we make the following assumption on the cost function.

Assumption 4: There exists a constant C such that, for all x, a, f_1, f_2 and m, we have

$$|c(x,a,f_1,m) - c(x,a,f_2,m)| \\ \leq C\,|c(x,a,f_1,m)|\,\|f_1 - f_2\|_{1-g} + H\left(\|f_1 - f_2\|_{1-g}^2\right)$$

Note that here $H(\|f\|_{1-g}^2)$ denotes a function that is linear in moments of $\|f\|$ up to the second moment; i.e., $H(\|f\|^2)$ cannot include any terms that depend on moments higher than $\|f\|_{1-g}^2$. Note that the function g in Definition 4 must be chosen such that the assumptions 3 and 4 are simultaneously satisfied. As we will show in section V, the linear quadratic Gaussian (LQG) tracking problem discussed in [2] is a special case where the difference in the cost functions can be expressed in the form given above. Furthermore,we will show that for the LQG problem, the function $g(y)$ is a polynomial in the state y.

The next assumption is on the set of policies $\mathcal{M}$. We will restrict attention to sequences $(\mu^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)})$ that satisfy the following assumption.

Assumption 5: Given a sequence $(\mu^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)})$, we assume that:

$$\sup_m \mathbb{E}\Big[\sum_{t=0}^{\infty} \beta^t c^2\left(x_{i,t}, a_{i,t}, \tilde{f}_{-i}^{(m)}, m\right) \ | \\ x_{i,0} = x, f_{-i,0}^{(m)} = f; \mu^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)}\Big] < \infty$$

In contrast to earlier work on oblivious equilibrium [1] [6], our assumptions are significantly different. Specifically, we do not assume that the cost functions are uniformly bounded in state x, distribution f, or actions a as was done in the previous work. This lack of uniform bound on the cost function necessitates a significantly different proof technique. However, some of the ideas in the proofs are borrowed from [6].

IV. ASYMPTOTIC RESULTS FOR OBLIVIOUS EQUILIBRIUM

In this section, we prove the AME property using a series of technical lemmas. Assumptions 1-5 are kept throughout the remainder of this section. The first lemma shows that under the $1-g$ norm, the variance of the distribution $f_{-i,t}^{(m)}$ goes to zero as the number of players become large.

Lemma 1: Under the light-tail assumption and if all the players use oblivious policy $\tilde{\boldsymbol{\mu}}_{-i}^{(m)}$, we have

$$\mathbb{E}\left[\left\|f_{-i,t}^{(m)} - \tilde{f}_{-i}^{(m)}\right\|_{1-g}^2\right] \to 0 \quad \text{as} \quad m \to \infty$$

Proof: We can write

$$\left\|f_{-i,t}^{(m)}(y) - \tilde{f}_{-i}^{(m)}(y)\right\|_{1-g} = \sum_y g(y)\left|f_{-i,t}^{(m)}(y) - \tilde{f}_{-i}^{(m)}(y)\right|.$$

Now, let a small $\epsilon > 0$ be given and let z be such that the light tail condition in equation (7) and (8) is satisfied for the given ϵ. Then,

$$\left\|f_{-i,t}^{(m)} - \tilde{f}_{-i}^{(m)}\right\|_{1,g}^2 \leq \Big[z \max_{|y|\leq z} g(y)\left|f_{-i,t}^{(m)} - \tilde{f}_{-i}^{(m)}(y)\right| \\ + \sum_{|y|>z} g(y) f_{-i,t}^{(m)}(y) + \sum_{|y|>z} g(y)\tilde{f}_{-i}^{(m)}(y)\Big]^2.$$

Using the identity $(a+b+c)^2 \leq 4a^2 + 4b^2 + 4c^2$, we get that

$$\mathbb{E}\left\|f_{-i,t}^{(m)} - \tilde{f}_{-i}^{(m)}\right\|_{1,g}^2 \leq \underbrace{4z^2\mathbb{E}\left[\max_{|y|\leq z} g^2(y)\left|f_{-i,t}^{(m)} - \tilde{f}_{-i}^{(m)}(y)\right|^2\right]}_{\equiv A_z^{(m)}} \\ + \underbrace{4\mathbb{E}\left[\sum_{|y|>z} g(y) f_{-i,t}^{(m)}(y)\right]^2}_{\equiv B_z^{(m)}} + \underbrace{4\mathbb{E}\left[\sum_{|y|>z} g(y)\tilde{f}_{-i}^{(m)}(y)\right]^2}_{\equiv C_z^{(m)}}. \tag{9}$$

Note that term in parenthesis in $C_z^{(m)}$ is independent of t and hence a constant. By the light tail assumption, for sufficiently small $\epsilon > 0$ and sufficiently large z, we have $C_z^{(m)} \leq 4\epsilon^2 < 4\epsilon$ for all m. Let us now consider the term $A_z^{(m)}$. We have

$$\begin{aligned} &\mathbb{E}\Big(f_{-i,t}^{(m)}(y) - \tilde{f}_{-i}^{(m)}\Big)^2 \\ &= \frac{1}{(m-1)^2}\mathbb{E}\left(\sum_{j\neq i} \mathbf{1}_{\{x_{j,t}=y\}} - \mathbb{E}\left(\sum_{j\neq i}\mathbf{1}_{\{x_{j,t}=y\}}\right)\right)^2, \\ &= \frac{1}{(m-1)^2}\sum_{j\neq i} \text{Var}\left(\mathbf{1}_{\{x_{j,t}=y\}}\right), \\ &\leq \frac{1}{4(m-1)} \to 0 \text{ as m} \to \infty. \end{aligned}$$

The random variable $\mathbf{1}_{\{x_{j,t}=y\}}$ is a Bernoulli random variable with $\mathbb{E}\left[\mathbf{1}_{\{x_{j,t}=y\}}\right] = q^{(m)}(y)$ and $\text{Var}\left(\mathbf{1}_{\{x_{j,t}=y\}}\right) = q^m(y)(1-q^m(y)) \leq \frac{1}{4}$. Thus, for each $|y| \leq z$, there exists an m_y such that for $m \geq m_y$ we have

$$\mathbb{E}\left[\left|f_{-i,t}^{(m)}(y) - \tilde{f}_{-i}^{(m)}(y)\right|^2\right] \leq \frac{\epsilon}{4z^2 g^2(y)}$$

Define $m_a = \max_{|y|\leq z}\{m_y\}$. Then, for $m > m_a$ we have $A_z^{(m)} < \epsilon$.

Let us now consider the term $B_z^{(m)}$. Using the definition of $f_{-i,t}^{(m)}$ we have

$$\begin{aligned} B_z^{(m)} &= \frac{4}{(m-1)^2}\mathbb{E}\left[\sum_{|y|>z}\sum_{j\neq i} g(y)\mathbf{1}_{\{x_{j,t}=y\}}\right]^2 \\ &= \frac{4}{(m-1)^2}\mathbb{E}\left[\sum_{j\neq i}\sum_{|y|>z} g(y)\mathbf{1}_{\{x_{j,t}=y\}}\right]^2 \end{aligned}$$

where the interchange of summation is justified since for any given m, the summations are finite. Let us denote

$\sum_{|y|>z} g(y)\mathbf{1}_{\{x_{j,t}=y\}} = W_{j,t}^{(m)}$. We can thus write $B_z^{(m)}$ as

$$B_z^{(m)} = \frac{4}{(m-1)^2}\mathbb{E}\left[\sum_{j\neq i} W_{j,t}^{(m)}\right]^2 = \frac{4}{(m-1)^2}\left[\mathrm{Var}\Big(\sum_{j\neq i} W_{j,t}^{(m)}\Big) + \Big(\sum_{j\neq i}\mathbb{E}W_{j,t}^{(m)}\Big)^2\right] \quad (10)$$

From equation (4), we know that $\mathbb{E}\left[\mathbf{1}_{\{x_{j,t}=y\}}\right] = \tilde{f}_{-i}^{(m)}$. From the light tail condition, for chosen z, we have $\mathbb{E}\Big(W_{j,t}^{(m)}\Big) < \epsilon$ for all m. Thus,

$$\sum_{j\neq i}\mathbb{E}\left(W_{j,t}^{(m)}\right) < (m-1)\epsilon \quad (11)$$

Let us now consider the first term in equation (10). We have

$$\begin{aligned}\mathrm{Var}\Big(\sum_{j\neq i} W_{j,t}^{(m)}\Big) &= \sum_{j\neq i}\mathrm{Var}\left(W_{j,t}^{(m)}\right)\\ &= (m-1)\mathrm{Var}\left(W_{j,t}^{(m)}\right)\\ &= (m-1)\left[\mathbb{E}\left(W_{j,t}^{(m)}\right)^2 - \left(\mathbb{E}\left(W_{j,t}^{(m)}\right)\right)^2\right]\\ &< (m-1)\mathbb{E}\left(W_{j,t}^{(m)}\right)^2 + (m-1)\epsilon^2\end{aligned} \quad (12)$$

where the first equality follows from the fact that $f_{-i,t}^{(m)}$ is an $(m-1)$ fold convolution of $\tilde{f}_{-i}^{(m)}$, since all the opponents are using oblivious policy $\tilde{\mu}$ and hence are evolving independently. The last inequality is because of the chosen value of z, which gives $\mathbb{E}\Big(W_{j,t}^{(m)}\Big) < \epsilon$ for all m. Note that $W_{j,t}^{(m)}$ as defined is a random variable that takes the value $g(y)$ with probability $\tilde{f}_{-i}^{(m)}(y)$. Thus,

$$\mathbb{E}\left(W_{j,t}^{(m)}\right)^2 = \sum_{|y|>z} g^2(y)\tilde{f}_{-i}^{(m)}(y) < \epsilon$$

where the last inequality follows from the light tail condition. Substituting the above equation in equation (12), we get that

$$\mathrm{Var}\Big(\sum_{j\neq i} W_{j,t}^{(m)}\Big) < (m-1)\epsilon + (m-1)\epsilon^2 < 2(m-1)\epsilon \quad (13)$$

Substituting equations (11) and (13) in equation (10), we get that for some $m > m_b$ we have

$$\begin{aligned}B_z^{(m)} &< \frac{4}{(m-1)^2}\left[2(m-1)\epsilon + (m-1)^2\epsilon^2\right]\\ &< \frac{8}{(m-1)}\epsilon + 4\epsilon^2 \ \le 12\epsilon\end{aligned}$$

From bounds on $A_z^{(m)}$ and $C_z^{(m)}$ and the above equation, we get that for sufficiently large z and for $m > \max\{m_a, m_b\}$ and we have

$$\mathbb{E}\left\|f_{-i,t}^{(m)} - \tilde{f}_{-i}^{(m)}\right\|_{1,g}^2 \le 17\epsilon.$$

Since ϵ is arbitrary, this proves the lemma. ■

Lemma 2: For all x and $(\mu^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)})$ satisfying assumptions 1-5, we have

$$\lim_{m\to\infty}\mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t\left|c\left(x_{i,t}, a_{i,t}, f_{-i,t}^{(m)}, m\right) - c\left(x_{i,t}, a_{i,t}, \tilde{f}_{-i}^{(m)}, m\right)\right| \mid x_{i,0} = x; \mu^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)}\right] = 0.$$

Proof: Let us define

$$\Delta_{i,t}^m \triangleq \left|c\left(x_{i,t}, a_{i,t}, f_{-i,t}^{(m)}, m\right) - c\left(x_{i,t}, a_{i,t}, \tilde{f}_{-i}^{(m)}, m\right)\right|.$$

Also denote $c_{m,t} = \left|c\left(x_{i,t}, a_{i,t}, \tilde{f}_{-i}^{(m)}, m\right)\right|$ and $F_{m,t} = \left\|f_{-i,t}^{(m)} - \tilde{f}_{-i}^{(m)}\right\|_{1-g}$. Using assumption 4 on the cost function we have

$$\begin{aligned}&\mathbb{E}\Big[\sum_{t=0}^{\infty}\beta^t\Delta_{i,t}^m \mid \mu^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)}\Big]\\ &\le C\mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t c_{m,t}F_{m,t} \mid \mu^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)}\right] + \mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t H\left(F_{m,t}^2\right)\right]\\ &= C\sum_{t=0}^{\infty}\beta^t\mathbb{E}\left[c_{m,t}F_{m,t} \mid \mu^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)}\right] + \sum_{t=0}^{\infty}\beta^t H\left(\mathbb{E}\left[F_{m,t}^2\right]\right)\end{aligned}$$

where the last equality follows from monotone convergence theorem. Let us denote the first term of the above equation as T_1 and the second term as T_2. Using Cauchy-Schwarz inequality we get that

$$\begin{aligned}T_1 &\le C\sum_{t=0}^{\infty}\left(\beta^t\mathbb{E}\left[c_{m,t}^2\right]\right)^{1/2}\left(\beta^t\mathbb{E}\left[F_{m,t}^2\right]\right)^{1/2}\\ &\le C\left(\sum_{t=0}^{\infty}\beta^t\mathbb{E}\left[c_{m,t}^2\right]\right)^{1/2}\left(\sum_{t=0}^{\infty}\beta^t\mathbb{E}\left[F_{m,t}^2\right]\right)^{1/2}\end{aligned}$$

where the last inequality is due to Hölder's inequality. Note that we have dropped the policy vector in the conditioning field for notational compactness. By assumption 5, the first term in above equation is bounded. Also, note that $\mathbb{E}\left[F_{m,t}^2\right]$ is independent of t and hence

$$\sum_{t=0}^{\infty}\beta^t\mathbb{E}\left[F_{m,t}^2\right] = \frac{\mathbb{E}\left[F_{m,t}^2\right]}{1-\beta}$$

Substituting the above equation in second term of T_1 and also in T_2 and using lemma 1 we get the desired result. ■

Theorem 3 (Main Theorem): Consider a sequence of oblivious equilibrium policies $\boldsymbol{\mu}^{(m)}$ that satisfies assumption 5 with $\mu^{(m)}$ equal to either the oblivious or non-oblivious best response to $\tilde{\boldsymbol{\mu}}_{-i}^{(m)}$. Then the AME property holds. That is, for all i, x, we have

$$\lim_{m\to\infty}\mathbb{E}\Big[V\left(x, f, m \mid \tilde{\mu}^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)}\right) - \inf_{\mu'\in\mathcal{M}} V\left(x, f, m \mid \mu', \tilde{\boldsymbol{\mu}}_{-i}^{(m)}\right)\Big] = 0.$$

Proof: The proof of the theorem is similar to one given in [1] and is omitted due to space constraints. ■

V. Linear Quadratic System

In this section, we consider linear quadratic (LQ) games with many players [2]. We show that these games are a special case of the model considered in the previous section. We keep assumptions 1, 2 and 5 throughout the remainder of this section. Our objective in this section is to prove that assumptions 3 and 4 hold for the LQ model, and thus the AME property holds for the LQ model. Compared to [2], we consider a discrete time version of the LQ games. Furthermore, for simplicity we assume that all players have the same form of cost function, i.e., they are uniform. This assumption can be relaxed, as we discuss in section VI. The state evolution of a player $1 \leq i \leq m$ is given as

$$x_{i,t+1} = Ax_{i,t} + Ba_{i,t} + w_{i,t}, \tag{14}$$

where $x_{i,t} \in \mathbb{Z}$ and $a_{i,t} \in \mathbb{Z}$. Here $A, B > 0$. The noise process $w_{i,t}$ is assumed to be independent across time as well as players. We also assume that the noise process has zero mean and all its moments are finite.

We assume that the single period cost function for a player i is separable in its action. That is

$$c\left(x_{i,t}, a_{i,t}, f_{-i,t}^{(m)}\right) = c_1\left(x_{i,t}, f_{-i,t}^{(m)}\right) + c_2\left(a_{i,t}\right)$$

where we assume that $c_2(a_{i,t})$ is a quadratic function of $a_{i,t}$. Note that the cost function does not depend upon the number of players playing the game. Here $f_{-i,t}^{(m)}$ is the distribution of m players over the state space. Let us define the kth moment of $f_{-i,t}^{(m)}$ as

$$\gamma_k = \sum_y f_{-i,t}^{(m)}(y)y^k.$$

We assume that c_1 is a jointly strictly convex, quadratic function of x and $\gamma_k,\ k = 1, \ldots, K$. Also, for given x and $f_{-i,t}^{(m)}$, we have

$$\inf_{x,f} c_1(x, f) = \epsilon_0 > 0.$$

This model is a generalization of the LQG games considered in [2], where the cost function depends only on the first moment γ_1 of the distribution $f_{-i,t}^{(m)}$. Furthermore, that result was derived for a specific form of cost function.

Each player choses an action $a_{i,t} = \mu^{(m)}\left(x_{i,t}, f_{-i,t}^{(m)}\right)$ to minimize its total expected cost. The expected net present cost for a player i is given as

$$V\left(x, f \mid \mu^{(m)}, \boldsymbol{\mu}_{-i}\right) =$$
$$\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t c\left(x_{i,t}, a_{i,t}, f_{-i,t}^{(m)}\right) \ | x_{i,0} = x, f_{-i,0}^{(m)} = f; \mu^{(m)}, \boldsymbol{\mu}_{-i}\right]$$

We assume that each player chooses its policy $\mu^{(m)}$ to make its expected net present value $V\left(x, f \mid \mu^{(m)}, \boldsymbol{\mu}_{-i}\right)$ finite.

In oblivious equilibrium, each player's policy is only a function of its current state and average aggregate distribution of its competitors. The next lemma shows that for the LQ model described above, the oblivious equilibrium is independent of the number of players m.

Lemma 4: Under assumption 2, for the LQ model described above, there exists an oblivious equilibrium that is independent of the number of players m playing the game.

Proof: Let $(\hat{\mu}, \hat{f})$ be an oblivious equilibrium for the LQ model with m_1 players playing the game. This means that $\hat{\mu}$ is an optimal policy under the cost function $c(x, a, \hat{f})$ and that $\hat{f}$ is the stationary distribution obtained from the dynamic equation $x_{i,t+1} = Ax_{i,t} + B\hat{\mu}(x_{i,t}, \hat{f}) + w_{i,t}$. Now let us assume that the number of players changes to m_2. Consider a player i and assume that every one of its $m_2 - 1$ opponent uses the policy $\hat{\mu}$. Then $\tilde{f}_{-i} = \hat{f}$, and since the cost function does not depend upon the total number of players, the optimal policy for player i is $\hat{\mu}$. Since the dynamics are independent of the number of players, the stationary distribution for player i is $\hat{f}$. Thus, $(\hat{\mu}, \hat{f})$ is an oblivious equilibrium for the game with m_2 players. ■

To prove that there is no optimality loss if a player uses oblivious equilibrium policy, we first define the weight function g for the LQ model.

Definition 5 (LQ 1-g norm): For the LQ model described above, let $g(y) = |y|^K$. With this weight function, we define

$$\|f\|_{1-g} = \sum_y f(y)|y|^K$$

The following lemma verifies the light-tail condition for the LQ model.

Lemma 5: For the LQ model, the light tail condition holds for the weight function $g(y) = |y|^K$.

Proof: Note that $\tilde{f}_{-i} = q(y)$ where $q(y)$ is the stationary probability of the Markov chain to be in state y and is independent of the number of players m. For linear systems with quadratic cost, we know that the optimal policy is a linear function of the state [11]. That is, $\tilde{\mu} = -Lx + l_0$, where $L > 0$ and l_0 are constants that may depend on $\tilde{f}$. The closed loop system thus evolves as $x_{t+1} = (A - BL)x_t + Bl_0 + w_t$. From assumption 1, we know that the closed loop system is stable; thus we have $(A - BL) < 1$.

We use the Foster-Lyapunov stability criterion [12] to establish the light tail condition. Using the Lyapunov function $V(x) = x$ and the fact that the noise process has finite mean, we can show that $\mathbb{E}_q(x)$ is finite. We then use induction to show that the stationary distribution has finite $2K$ moments. Specifically, assume that $\mathbb{E}_q(x^j)$ is finite for $j = 1, \ldots, p$, where $p < 2K$. Then, using the Lyapunov function $V(x) = x^{p+1}$ and the fact that the noise process has finite moments, we can show that $\mathbb{E}_q(x^{p+1})$ is finite. Thus, the stationary distribution has its first $2K$ moments finite [13]. So for any given ϵ, there exists a state z, such that

$$\sum_{|y|>z} |y|^{2K} q(y) < \epsilon.$$

Hence the lemma is proved. ■

The next lemma relates the absolute difference in the cost function to the $1 - g$ norm difference between the actual distribution $f_{-i,t}^{(m)}$ and its average $\tilde{f}_{-i}$.

Lemma 6: For the LQ model described above, there exist constants M_1, M_2 and M_3 such that

$$\left|c\left(x_{i,t}, a_{i,t}, f_{-i,t}^{(m)}\right) - c\left(x_{i,t}, a_{i,t}, \tilde{f}_{-i}\right)\right| \leq M_1 c\left(x_{i,t}, a_{i,t}, \tilde{f}_{-i}\right)\left\|f_{-i,t}^{(m)} - \tilde{f}_{-i}\right\|_{1-g} + M_2\left\|f_{-i,t}^{(m)} - \tilde{f}_{-i}\right\|_{1-g} + M_3\left\|f_{-i,t}^{(m)} - \tilde{f}_{-i}\right\|_{1-g}^2$$

Proof: For simplicity of notation, we drop the subscripts i,t. Also denote $\mathbf{\Gamma}(f) = [\gamma_1(f), \ldots, \gamma_K(f)]$, where $\gamma_i(f)$ is the ith moment of the distribution f. Then, we have

$$\begin{aligned}\Delta c &= \left|c\left(x, a, \mathbf{\Gamma}(f)\right) - c\left(x, a, \mathbf{\Gamma}(\tilde{f})\right)\right| \\ &= \left|c_1\left(x, \mathbf{\Gamma}(f)\right) - c_1\left(x, \mathbf{\Gamma}(\tilde{f})\right)\right| \\ &= \left|\int_{\alpha=0}^{1} \frac{\partial c_1\left(x, \mathbf{\Gamma}(\tilde{f} + \alpha(f - \tilde{f}))\right)}{\partial \alpha} d\alpha\right| \\ &\leq \int_{\alpha=0}^{1} \sum_y \left|\tilde{f}(y) - f(y)\right| \left|\frac{\partial c_1\left(x, \mathbf{\Gamma}(\tilde{f} + \alpha(f - \tilde{f}))\right)}{df(y)}\right| d\alpha \qquad (15)\end{aligned}$$

where we have used the fundamental theorem of calculus. Consider the second term in the above equation. We have

$$\begin{aligned}\left|\frac{\partial c_1\left(x, \mathbf{\Gamma}(\tilde{f} + \alpha(f - \tilde{f}))\right)}{df(y)}\right| &\leq \sum_{j=1}^{K}\left|\frac{\partial c_1}{\partial \gamma_j}\frac{\partial \gamma_j}{\partial f(y)}\right| \\ &= \sum_{j=1}^{K}\left|\frac{\partial c_1}{\partial \gamma_j}\right| |y|^j \\ &\leq K_1 |y|^K \sum_{j=1}^{K}\left|\frac{\partial c_1}{\partial \gamma_j}\right|\end{aligned}$$

where the last inequality holds for some K_1 since $y \in \mathbb{Z}$. Substituting the above equation in equation (15), we get

$$\Delta c \leq K_1\left\|\tilde{f} - f\right\|_{1-g} \sum_{j=1}^{K}\int_{\alpha=0}^{1}\left|\frac{\partial c_1\left(x, \mathbf{\Gamma}(\tilde{f} + \alpha(f - \tilde{f}))\right)}{\partial \gamma_j}\right| d\alpha \qquad (16)$$

where we use the fact that $\left\|\tilde{f} - f\right\|_{1-g}$ is independent of α and we also changed the order of integral and the summation.

Define a vector $z = [\gamma_1, \gamma_2, \ldots \gamma_K, x, 1]$ and note that c_1 is a strictly convex quadratic function of z. Thus, we can write $c_1(x, \gamma_1, \ldots, \gamma_K) = z^T Q z$ for some positive definite symmetric matrix Q. Note that $\mathbf{\Gamma}(\tilde{f} + \alpha(f - \tilde{f}))_j = \tilde{\gamma}_j + \alpha(\gamma_j - \tilde{\gamma}_j)$. Since c_1 is a quadratic function of γ_j, the partial derivative of c_1 with respect to γ_j is a linear function of γ_j and consequently a linear function of α. Thus, we have

$$\begin{aligned}&\left|\frac{\partial c_1\left(x, \mathbf{\Gamma}(\tilde{f} + \alpha(f - \tilde{f}))\right)}{\partial \gamma_j}\right| \\ &= \left|\sum_{s=1}^{K} Q_{sj} z_s + \sum_{s=1}^{K} Q_{js} z_s + Q_{K+1,j}x + 2Q_{j,K+2}\right| \\ &\leq \left|Q_{K+1,j}x\right| + \sum_{s=1}^{K}\left(|Q_{sj}| + |Q_{js}|\right)|z_s| + 2\left|Q_{j,K+2}\right| \\ &\leq \left|Q_{K+1,j}x\right| + \sum_{s=1}^{K}\left(|Q_{sj}| + |Q_{js}|\right)\left(|\tilde{\gamma}_s| + \alpha\left|\gamma_s - \tilde{\gamma}_s\right|\right) \\ &\quad + 2\left|Q_{j,K+2}\right| \qquad (17)\end{aligned}$$

where we have used the fact that $z_s = \tilde{\gamma}_s + \alpha(\gamma_s - \tilde{\gamma}_s)$ for $s = 1, \ldots, K$. Now,

$$\begin{aligned}|\gamma_s - \tilde{\gamma}_s| &= \left|\sum_y f(y) y^s - \sum_y \tilde{f}(y) y^s\right| \\ &\leq \sum_y \left|f(y) - \tilde{f}(y)\right| |y|^s \\ &\leq \sum_y \left|f(y) - \tilde{f}(y)\right| |y|^K \\ &= \left\|f - \tilde{f}\right\|_{1-g}\end{aligned}$$

Substituting above equation in equation (17), and integrating with respect to α we get that

$$\int_{\alpha=0}^{1}\left|\frac{\partial c_1\left(x, \mathbf{\Gamma}(\tilde{f} + \alpha(f - \tilde{f}))\right)}{\partial \gamma_j}\right| d\alpha \leq D_j^T |\tilde{z}| + K_2\left\|f - \tilde{f}\right\|_{1-g},$$

for some K_2. Here we used the notation $|\tilde{z}| = [|\tilde{\gamma}_1|, \ldots |\tilde{\gamma}_K|, |x|, 1]$ and D_j is a vector of length $K+2$ with coefficients from equation (17). Substituting above equation in equation (16), we get

$$\begin{aligned}\Delta c &\leq K_1\left\|\tilde{f} - f\right\|_{1-g}\left(\sum_{j=1}^{K} D_j^T |\tilde{z}| + K_2 K\left\|f - \tilde{f}\right\|_{1-g}\right) \\ &= K_1\left\|\tilde{f} - f\right\|_{1-g} D^T |\tilde{z}| + K_1 K_2 K\left\|f - \tilde{f}\right\|_{1-g}^2 \qquad (18)\end{aligned}$$

Now consider $D^T|\tilde{z}|$. We show that there exists a constant $\delta > 0$ such that $\delta D^T |\tilde{z}| \leq |\tilde{z}|^T Q |\tilde{z}| + \epsilon_0$. Define $w = (\delta/2)(Q)^{-1/2} D$. Then, we have

$$\begin{aligned}&|\tilde{z}|^T (Q) |\tilde{z}| + \epsilon_0 - \delta D^T |\tilde{z}| \\ &= |\tilde{z}|^T (Q) |\tilde{z}| - 2w^T (Q)^{1/2} |\tilde{z}| + \epsilon_0 + w^T w - w^T w \\ &= \left((Q)^{1/2}|\tilde{z}| - w\right)^T \left((Q)^{1/2}|\tilde{z}| - w\right) + \epsilon_0 - w^T w \\ &= \left(|\tilde{z}| - (Q)^{-1/2} w\right)^T Q \left(|\tilde{z}| - (Q)^{-1/2} w\right) + \epsilon_0 - w^T w \\ &\geq \epsilon_0 - w^T w\end{aligned}$$

where the last inequality follows since the first term is a quadratic form. So we need to ensure that $\epsilon_0 - w^T w \geq 0$, which implies that

$$\frac{\delta^2}{4} D^T (Q)^{-1} D \leq \epsilon_0 \quad \Longrightarrow \quad \delta \leq \sqrt{\frac{4\epsilon_0}{D^T (Q)^{-1} D}}$$

If we choose $K_3 > 1/\delta$, then we have $D^T|\tilde{z}| \leq K_3 |\tilde{z}|^T Q |\tilde{z}| + K_3 \epsilon_0$.

Next we show that there exists a K_4 such that $|\tilde{z}|^T Q |\tilde{z}| \leq K_4 \tilde{z} Q \tilde{z}$, where $\tilde{z} = [\tilde{\gamma}_1, \ldots \tilde{\gamma}_K, x, 1]$. Note that $|\tilde{z}|^T |\tilde{z}| = \tilde{z}^T \tilde{z}$. We have

$$\begin{aligned} |\tilde{z}|^T Q |\tilde{z}| &\leq \lambda_{\max}^Q |\tilde{z}|^T |\tilde{z}| \\ &= \lambda_{\max}^Q \tilde{z}^T \tilde{z} \\ &\leq K_4 \lambda_{\min}^Q \tilde{z}^T \tilde{z} \end{aligned}$$

where the last inequality is true for a suitable choice of K_4. Note that Q is positive definite matrix so $\lambda_{\min}^Q > 0$. Since $K_4 \lambda_{\min}^Q \tilde{z}^T \tilde{z} \leq K_4 \tilde{z}^T Q \tilde{z}$, we have that $|\tilde{z}|^T Q |\tilde{z}| \leq K_4 \tilde{z} Q \tilde{z}$ for some value of K_4. Now,

$$\begin{aligned} D^T |\tilde{z}| &\leq K_3 |\tilde{z}|^T Q |\tilde{z}| + K_3 \epsilon_0 \\ &\leq K_3 K_4 \tilde{z} Q \tilde{z} + K_3 \epsilon_0 \\ &= K_3 K_4 c_1 \left(x, \tilde{\gamma_1}, \ldots \tilde{\gamma_K}\right) + K_3 \epsilon_0 \end{aligned} \tag{19}$$

Substituting the above equation in equation (18), we prove the lemma. ■

We restrict attention to sequences $(\mu^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)})$ that satisfy assumption 5. We conjecture that this assumption would be satisfied for the LQ model if the noise process has finite fourth moment. Under this assumption, the AME property holds for linear systems with quadratic costs.

VI. CONCLUSIONS AND DISCUSSION

In this paper, we studied stochastic dynamic games with many players, where the players are coupled via their cost functions. Similar to [1], we showed that for a certain class of stochastic games, we can approximate MPE by a computationally simpler concept called OE. Previous work done in this area [1], [6] (where the AME property was established for profit maximization) assumed a uniform bound on the cost function. In this paper, we extended the notion of OE to a class of games where the cost function is unbounded. The lack of a uniform bound necessitated a new proof technique. We showed that games with linear dynamics and quadratic costs are a special case of the model considered. This generalizes a similar result derived in [2].

In the development throughout the paper, we have assumed that all players are uniform, i.e., they have same form of cost function c. The results of this paper can be easily extended to the case where the players are non-uniform and their cost functions are drawn from a finite set of types. The reader is referred to [6], where a similar development was done albeit for a different set of assumptions. However, the same technique can be used here to extend our model; the details of which are omitted due to space constraints.

As mentioned before, the concept of oblivious equilibrium was first introduced in [1], where it was used to establish AME property for industry dynamic models. Common single period profit functions used in those models are bounded over states. A typical example is a profit function arising from price competition among firms that face a logit demand system generated by consumers with bounded income. For such functions, the AME property can be established using a slightly different approach. Specifically, assumptions 4 and 5 can be replaced by a uniform bound on the cost function as well as assuming that the cost functions are Gateaux differentiable with respect to $f^{(m)}(y)$. For such models, it is possible to show that

$$|\log c(x, a, f_1, m) - \log c(x, a, f_2, m)| \leq \|f_1 - f_2\|_{1-g}$$

where

$$g(y) = \sup_{x, a, f^{(m)}, m} \left| \frac{\partial \log c\left(x, a, f^{(m)}, m\right)}{\partial f^{(m)}(y)} \right|$$

The statement of lemma 2 then follows by a similar argument as given in lemma 4 of [6]. As a part of our future work, we hope to develop a model that unifies these two approaches.

REFERENCES

[1] G. Y. Weintraub, C. L. Benkard, and B. Van Roy, "Markov Perfect Industry Dynamics with Many Firms," accepted for publication, *Econometrica*.

[2] M. Huang, P. E. Caines and R. P. Malhamé, "Large-Population Cost-Coupled LQG Problems With Nonuniform Agents : Individual-Mass Behavior and Decentralized ϵ-Nash Equilibria," *IEEE Transactions on Automatic Control*, Vol. 52, No. 9, pp. 1560–1571, 2007.

[3] M. Huang, P. E. Caines and R. P. Malhamé, "Uplink Power Adjustment in Wireless Communication Systems: A Stochastic Control Analysis," *IEEE Transactions on Automatic Control*, Vol. 49, No. 10, pp. 1693-1708, October 2004.

[4] S. Ulukus and R. D. Yates, "Stochastic Power Control for Cellular Radio Systems," *IEEE Transactions on Communications*, Vol. 46, No. 6, pp. 784-798, 1998.

[5] R. Ericson and A. Pakes, "Markov-perfect industry dynamics: A framework for empirical work," *Review of Economic Studies* 62(1), pp. 53 - 82, 1995.

[6] V. Abhishek, S. Adlakha, R. Johari and G. Weintraub, "Oblivious Equilibrium for General Stochastic Games with Many Players," *Allerton Conference on Communication, Control and Computing*, 2007.

[7] A. Pakes and P. McGuire, "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model," *RAND Journal of Economics*, 25(4), pp. 555 589, 1994.

[8] A. Pakes and P. McGuire, "Stochastic Algorithms, Symmetric Markov Perfect Equilibrium, and the Curse of Dimensionality," *Econometrica*, 69(5), pp. 1261 - 1281, 2001.

[9] A. C. Antoulas and D. C. Sorensen, "Approximation of Large-Scale Dynamical Systems: An Overview," *Int. J. Appl. Math. Comput. Sci*, Vol. 11, No. 5, pp. 1093-1121, 2001.

[10] D. D. Siljak, "Decentralized Control of Complex Systems," *Academic Pr.*, 1991.

[11] D. P. Bertsekas, "Dynamic Programming and Optimal Control," *Athena Scientific*, 2000.

[12] S. P. Meyn and R. L. Tweedie, "Markov Chains and Stochastic Stability," *Springer-Verlag*, 1993.

[13] J. G. Dai and S. P. Meyn, "Stability and Convergence of moments for Multiclass Queuing Networks Via Fluid Limit Models," *IEEE Transactions on Automatic Control*, Vol. 40, No. 11, pp. 1889-1904, 1995.

[14] M. Huang, R. P. Malhamé, and P. E. Caines, "Nash Equilibria for Large-Population Linear Stochastic Systems of Weakly Coupled Agents," *Analysis, Control and Optimization of Complex Dynamical Systems*, pp. 215-252, 2005

[15] J. M. Lasry and P. L Lions, "Mean Field Games," *Japanese Journal of Mathematics*, 2, 229-260, 2007.