Oblivious Equilibrium for Stochastic Games with Concave Utility

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

***Abstract*— We study stochastic games with a large number of players, where players are coupled via their payoff 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 other players. This makes MPE computationally prohibitive as the number of players becomes large. An approximate solution concept called *oblivious equilibrium* (OE) was introduced by Weintraub et al., 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 that under a set of assumptions, as the number of players becomes large, OE closely approximates MPE. However, these assumptions require the computation of OE and verifying that the resulting stationary distribution satisfies a certain *light-tail* condition. In this paper, we derive exogenous conditions on the state dynamics and the payoff function under which the light-tail condition holds. A key condition is that the agents' payoffs are concave in their own state and actions. These exogenous conditions enable us to characterize a family of stochastic games in which OE is a good approximation for MPE.

I. INTRODUCTION

In this paper, we study stochastic games with large a number of players. Such games are used to model complex dynamical systems in both engineering [2], [3] and economics [4]. *Markov perfect equilibrium* is a commonly used equilibrium concept for stochastic games [8]. In MPE, strategies of players depend only on the current state of all players, and not on the past history of the game. In general, finding an MPE is analytically intractable; MPE is typically obtained numerically using dynamic programming (DP) algorithms [6]. As a result, the complexity associated with MPE computation increases rapidly with the number of players, the size of the state space, and the size of the action sets [7]. This limits its application to problems with small dimensions. Several techniques have been proposed in the literature to deal with the complexity of large scale systems [9], [10], [13], [14].

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 the computational difficulties associated with MPE: 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? In [1], the authors define a condition called the "light-tail" condition. The light-tail condition measures the impact of any deviation from the average behavior on the payoff of a player. Informally, this condition implies that the effect of a small perturbation in the instantaneous state of the competitors has a small effect on the payoff 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. As presented in [1] and the related exposition in [5], under a reasonable set of technical conditions (including the "light-tail" condition), OE is a good approximation to MPE.

The light-tail condition defined in [1] is an endogenous condition. One needs to compute the oblivious equilibrium to determine if OE is a good approximation to MPE. In this paper, we present exogenous conditions under which the light-tail condition holds. The exogenous conditions are on model primitives and hence are easy to verify. Thus, for a class of models satisfying the assumptions mentioned in this paper, we show that OE is a good approximation to MPE.

The rest of the paper is organized as follows. In Section II, we outline our model, describe our assumptions and define oblivious equilibrium. The structural assumptions required to prove the light-tail condition are also introduced in this section. In Section III, we introduce the asymptotic Markov equilibrium property and formally define the light tail condition. Section IV proves that under the assumptions made in the paper, the light-tail condition holds. Section V concludes the paper.

II. MODEL, DEFINITIONS AND ASSUMPTIONS

We consider an m-player stochastic game evolving over discrete time periods with an infinite horizon. The discrete time periods are indexed with $t \in \mathbb{Z}_+$, where $\mathbb{Z}_+$ is the set of non-negative integers. The state of player i at time t is denoted by $x_{i,t} \in \mathbb{Z}_+$. We assume that the state evolution for player i depends only on its current state and the action it takes. Specifically, we assume that the state dynamics are linear and are given as:

$$x_{i,t+1} = Ax_{i,t} + Ba_{i,t} + w_{i,t}, \qquad (1)$$

where $a_{i,t} \in \mathbb{Z}_+$ is the action taken by player i at time t. Here, $w_{i,t}$ is the noise process that is assumed to be

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

independent across players and across time. Furthermore, we assume for the remainder of the paper that *the noise process* $w_{i,t}$ *has finite support at each* t.

Assumption 1: We assume that $0 < A < 1$ and that $B > 0$.

The condition $A < 1$ ensures that in absence of any action, the state decays to zero and hence the system is stable. The assumption $B > 0$ ensures that any action "increases" the state of the system.

The single period payoff to player i at time t is given as $\pi(x_{i,t}, a_{i,t}, \boldsymbol{x}_{-i,t})$. Here $\boldsymbol{x}_{-i,t}$ is the state of all players except player i at time t. Note that the payoff to player i does not depend on the actions taken by other players. Furthermore, we assume that the payoff function is independent of the identity of other players. That is, it only depends on the current state $x_{i,t}$ of the player i, the action $a_{i,t}$ taken by the player i and for each y, the fraction $f_{-i,t}^{(m)}(y)$, which is the fraction of the players (excluding player i) that have their state as y. Thus, in a slight abuse of notation we can write the payoff to i at time t as $\pi\left(x_{i,t}, a_{i,t}, f_{-i,t}^{(m)}\right)$. Note that we can write $f_{-i,t}^{(m)}$ as

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

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

We make the following structural assumption on the single period payoff function $\pi(x, a, f)$.

Assumption 2: We assume that the payoff function is separable in state and action. That is

$$\pi(x, a, f) = \pi_1(x, f) + \pi_2(a),$$

where $\pi_1(x, f) \geq 0$ and $\pi_2(a) \leq 0$. We also assume that for given f, the payoff function is concave in state x and action a.

Thus, $\pi_2(a)$ represents the cost of action.

Let $\mu_i^{(m)}$ be the *policy* chosen by player i; that is, at time t player i chooses an action $a_{i,t} = \mu_i^{(m)}(x_{i,t}, f_{-i,t}^{(m)})$. 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 $V(x, f \mid \mu_i^{(m)}, \boldsymbol{\mu}_{-i}^{(m)})$ 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

$$V(x, f \mid \mu_i^{(m)}, \boldsymbol{\mu}_{-i}^{(m)}) \triangleq \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \pi(x_{i,t}, a_{i,t}, f_{-i,t}^{(m)}) \;\middle|\; x_{i,0} = x, f_{-i,0}^{(m)} = f; \mu_i^{(m)}, \boldsymbol{\mu}_{-i}^{(m)} \right], \quad (3)$$

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. Also note that the value function depends on the number of players m via the policy vector $\boldsymbol{\mu}^{(m)}$.

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. We make the following assumption on the set of policies $\mathcal{M}$.

Assumption 3: We restrict attention to a set of policies $\mathcal{M}$ such that for all x, there holds:

$$\sup_{\mu' \in \mathcal{M}} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \sup_f \pi_1(x_{i,t}, f) \mid x_{i,0} = x; \mu_i = \mu' \right] < \infty. \quad (4)$$

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

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

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 set of oblivious policies.

Assumption 4: We restrict attention to those oblivious policies that make each player's state Markov chain positive recurrent.

In a slight abuse of notation, we use $q^{(m)}$ to denote the stationary distribution associated with the oblivious policy $\tilde{\mu}^{(m)}$. The stationary distribution depends on the number of players m because the oblivious policy chosen by a player (may) depend 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 players except 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 distribution of its competitors is denoted by $\tilde{f}^{(m)}$, and is defined as:

$$\tilde{f}^{(m)}(y) \triangleq \mathbb{E}\left[f_{-i,t}^{(m)}(y) \mid \tilde{\boldsymbol{\mu}}_{-i}^{(m)}\right] = q^{(m)}(y) \tag{5}$$

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

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* $\tilde{V}(x \mid \tilde{\mu}^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)})$ 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

$$\tilde{V}(x \mid \tilde{\mu}^{(m)}, \tilde{\boldsymbol{\mu}}_{-i}^{(m)}) \triangleq \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \pi\left(x_{i,t}, a_{i,t}, \tilde{f}^{(m)}\right) \;\middle|\; x_{i,0} = x; \tilde{\mu}^{(m)} \tilde{\boldsymbol{\mu}}_{-i}^{(m)}\right]. \tag{6}$$

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 $\tilde{f}^{(m)}$. 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 $\tilde{\boldsymbol{\mu}} \in \tilde{\mathcal{M}}$ represents an *oblivious equilibrium* if for all i, we have

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

In this paper, we do not show the existence of Markov perfect equilibrium or of oblivious equilibrium.

Assumption 5: Markov perfect equilibrium and oblivious equilibrium exist for the stochastic game under consideration.

III. Asymptotic Markov Equilibrium and the Light-Tail Condition

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, as first defined by [1]. 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

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

Note that the expectation here is over $f^{(m)}$, which denotes the aggregate state of all players other than i. Here the initial states of the m players are drawn out of the oblivious stationary distribution $\tilde{f}^{(m)}$. MPE requires the error to be zero for all $(x, f^{(m)})$, 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 under the OE.

In order to establish the AME property, we make some assumptions on the payoff functions. For notational convenience, we drop the subscripts i, t whenever it does not lead to any ambiguity.

Assumption 6: We assume that $\log \pi_1$ is Gateaux differentiable with respect to $f^{(m)}(y)$. That is, if

$$\sum_y \Delta f^{(m)}(y) \left|\frac{\partial \log \pi_1(x, f^{(m)})}{\partial f^{(m)}(y)}\right| < \infty,$$

then

$$\left.\frac{\partial \log \pi_1\left(x, f^{(m)} + \gamma \Delta f^{(m)}\right)}{\partial \gamma}\right|_{\gamma=0} = \sum_y \Delta f^{(m)}(y) \frac{\partial \log \pi_1\left(x, f^{(m)}\right)}{\partial f^{(m)}(y)}.$$

We define $g(y)$ as

$$g(y) = \sup_{x, f^{(m)}} \left|\frac{\partial \log \pi_1(x, f^{(m)})}{\partial f^{(m)}(y)}\right| \tag{7}$$

and assume that $g(y)$ is finite for all y. The function $g(y)$ can be interpreted as the maximum rate of change of the logarithm of the single period payoff of any player, with respect to a small change in the fraction of competitors at any state value y.

Assumption 7: We assume that the payoff function $\pi(x, a, f)$ is such that $g(y) \in \mathcal{O}(y^k)$ for some k.

Thus, we assume that whenever the fraction of competitors at any state y changes, the payoff of a player changes at most as a monomial of order k. We now proceed to define the *light-tail condition* formally.

Definition 4 (Light-Tail): We say that a sequence of oblivious distributions $\tilde{f}^{(m)}$ has a *light tail* if for every $\epsilon > 0$ there exists a state value z, such that

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

Here $\tilde{U}^{(m)}$ is a random variable distributed according to $\tilde{f}^{(m)}$.

The light tail requires that the probability of competitors at larger y (tail probability) should go to zero quickly uniformly over m. The quantity $\mathbb{E}\left[g(\tilde{U}^{(m)})\mathbf{1}_{\tilde{U}^{(m)}>z|\tilde{U}^{(m)}} \sim \tilde{f}^{(m)}\right]$ captures the average effect of competitors at a higher state on the single period payoff of a player.

IV. Exogenous Conditions and the Light Tail

We now establish that under the assumptions on the payoff function, the light-tail condition is satisfied for all $g(y) \in \mathcal{O}(y^k)$ for some k. Note that 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 under the assumptions made above, the oblivious equilibrium is independent of the number of players m.

Lemma 1: Under the assumptions made 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 with m_1 players playing the game; recall that we have assumed at least one OE exists. This means that $\hat{\mu}$ is an optimal policy under the payoff function $\pi(x, a, \hat{f})$ 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 $m_2 - 1$ opponents uses the policy $\hat{\mu}$. Then $\tilde{f}_{-i} = \hat{f}$, and since the payoff function does not depend upon the total number of players, the optimal policy for player i is $\hat{\mu}$. Note that the dynamics are also independent of the number of players. Thus, the stationary distribution for player i is $\hat{f}$. Hence $(\hat{\mu}, \hat{f})$ is an oblivious equilibrium for the game with m_2 players. ∎

In the light of above lemma, the light tail condition simplifies: for all $\epsilon > 0$, there must exist a z such that:

$$\mathbb{E}\left[g(U)\mathbf{1}_{\tilde{U}>z} \mid U \sim \tilde{f}\right] \leq \epsilon. \tag{9}$$

This condition is equivalent to the requirement that $g(U)$ *has finite expectation*.

Before we prove the light tail condition, we first show that under the structural assumptions made on the payoff function, the optimal policy is a decreasing function of the state x. We drop the subscript i for notational clarity.

Lemma 2: Under the assumptions made above, the optimal policy $\tilde{\mu}(x, \tilde{f})$ is a decreasing function of the state x.

Proof: From the principle of dynamic programming [11], we know that the optimal policy is given by Bellman's equation:

$$\tilde{\mu}(x, \tilde{f}) = \arg\max_a \left\{\pi(x, a, \tilde{f}) + \beta\mathbb{E}\left[V(Ax + Ba + w, \tilde{f})\right]\right\}, \tag{10}$$

where the expectation is taken over the statistics of the noise process w. Here $V(x, \tilde{f})$ is the value function and is given as

$$V(x, \tilde{f}) = \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \pi\left(x_t, a_t, \tilde{f}\right) \;\middle|\; x_0 = x; \tilde{\mu}\right].$$

Since the payoff function π is concave function of its state x and the action a, it can be shown that the resulting value function is also a concave function of x for every $\tilde{f}^{(m)}$ [15].

Let us define $J(x, a, \tilde{f})$ as

$$J(x, a, \tilde{f}) = \pi(x, a, \tilde{f}) + \beta\mathbb{E}\left[V(Ax + Ba + w, \tilde{f})\right],$$

and define $\Delta J(x, a, \tilde{f}) = J(x, a+1, \tilde{f}) - J(x, a, \tilde{f})$. It is easy to check that

$$\Delta J(x, a, \tilde{f}) = \pi_2(a+1) - \pi_2(a) + \beta\mathbb{E}\left[\Delta V(x, a, \tilde{f}, w)\right]$$

where $\Delta V(x, a, \tilde{f}, w) = V(Ax + B(a+1) + w, \tilde{f}) - V(Ax + Ba + w, \tilde{f})$. Denote $a^*(x, \tilde{f})$ to be the action that achieves the maximum in the equation (10). Then for $a < a^*$, $\Delta J(x, a, \tilde{f}) \geq 0$ and for $a \geq a^*$, $\Delta J(x, a, \tilde{f}) \leq 0$. In other words, $\Delta J(x, a, \tilde{f})$ changes sign at a^*.

Now let $x' > x$. Then $\Delta J(x', a, \tilde{f})$ is given as

$$\Delta J(x', a, \tilde{f}) = \pi_2(a+1) - \pi_2(a) + \beta\mathbb{E}\left[\Delta V(x', a, \tilde{f}, w)\right]$$

Note that $V(x, \tilde{f})$ is a concave function of x. Thus, $\Delta V(x, a, \tilde{f}, w)$ is a decreasing function of x. This gives that

$$\Delta J(x', a, \tilde{f}) \leq \Delta J(x, a, \tilde{f}).$$

Thus, the action a at which $\Delta J(x', a, \tilde{f})$ changes sign is smaller than the action at which $\Delta J(x, a, \tilde{f})$ changes sign. Hence, the optimal action $a^*(x, \tilde{f})$ decreases with x. ∎

Lemma 3: Under the assumptions made above, with $g(y) \in \mathcal{O}(y^k)$ for some k, there holds $\mathbb{E}[g(U)|U \sim \tilde{f}] < \infty$; i.e., the light-tail condition holds.

Proof: In order to prove the lemma, it suffices to show that the distribution $\tilde{f}$ has its first k moments finite. From the state equation (1), we see that the closed loop system evolves as

$$x_{t+1} = Ax_t + B\tilde{\mu}(x_t, \tilde{f}) + w_t, \tag{11}$$

where we have dropped the subscript i for brevity.

Choose a constant C such that $A < C < 1$. Then since $\tilde{\mu}(x, \tilde{f})$ is nonincreasing in x, we conclude that for all sufficiently large x we have:

$$\tilde{\mu}(x, \tilde{f}) \leq \frac{(C - A)}{B} x.$$

Rearranging terms, we find:

$$Ax + B\tilde{\mu}(x, \tilde{f}) \leq Cx$$

for all sufficiently large x. Using the Foster-Lyapunov criterion [12], it is straightforward to show that the closed loop system $x_{t+1} = Cx_t + w_t$ has k finite moments as long as the noise process w_t has k finite moments. Since we have assumed that w_t has finite support, the latter condition trivially holds. Thus the steady state distribution obtained from equation (1) has finite moments. This proves the lemma. ∎

V. Conclusion

Having proven the light-tail condition, the remainder of the proof follows closely the development of [1], as applied to general stochastic games in [5]. We thus omit the details.

ACKNOWLEDGMENTS

The authors gratefully acknowledge helpful conversations with Prof. Lanier Benkard and Prof. Benjamin Van Roy of Stanford University. This work was supported in part by the Stanford Clean Slate Internet Program, and by DARPA under the ITMANET program.

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, "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.

[3] 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.

[4] R. Ericson and A. Pakes, "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies* 62(1), pp. 53 - 82, 1995.

[5] 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.

[6] 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.

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

[8] L. S. Shapley, "Stochastic Games," *Proceeding of the National Academy of Sciences*, 39, pp. 1095-1100, 1953.

[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] 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

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

[15] J. E. Smith and K. F. McCardle, "Structural Properties of Stochastic Dynamic Programs, " *Operations Research*, Vol. 50, No. 5, pp. 796-809, 2002.