Oblivious Equilibrium: An Approximation to Large Population Dynamic 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 such games is Markov perfect equilibrium (MPE). It is well known that the computation of MPE suffers from the "curse of dimensionality." Recently an approximate solution concept called "oblivious equilibrium" (OE) was developed by Weintraub et. al, where each player reacts to only the average behavior of other players. In this work, we characterize a set of games in which OE approximates MPE. Specifically, we show that if system dynamics and payoff functions are concave in state and action and have decreasing differences in state and action, then an oblivious equilibrium of such a game approximates MPE. These exogenous conditions on model primitives allow us to characterize a set of games where OE can be used as an approximate solution concept.**

Extended Abstract

We study stochastic games with a large number of players. Such games are used to model complex dynamical systems in both engineering [1], [2] and economics [3]. A commonly used equilibrium concept for such games is that of *Markov perfect equilibrium (MPE)*. 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 algorithms [4]. 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 [5]. 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 [6], [7], [8], [9].

Recently, a scheme for approximating MPE for such large scale games was proposed in [10], via a solution concept called *oblivious equilibrium (OE)*. 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.

S. Adlakha and A. Goldsmith are with the Department of Electrical Engineering, Stanford University, 350 Serra Mall, Stanford, CA, USA. adlakha@stanford.edu, andrea@wsl.stanford.edu.

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

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

This work was supported by Stanford clean slate network design and DARPA ITMANET programs.

As proved in [10], the oblivious equilibrium approximates Markov perfect equilibrium under 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. The light-tail condition as defined in [10] is an endogenous condition: the OE must first be computed in order to verify the condition. In this work, we characterize a set of games under which OE approximates MPE.

Our main contributions are as follows:

1) We characterize a set of conditions on primitives of the model which ensure that OE approximates MPE. Specifically, we give conditions on the state dynamics and the payoff function which lead to the light-tail condition.
2) We also partially unify earlier related models, especially [11] and [10], in a single framework. In some respects, we also generalize the model given in both these prior works as well as our previous models in [12].

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 given as:

$$x_{i,t+1} = h\big(x_{i,t}, a_{i,t}, w_{i,t}\big), \tag{1}$$

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 independent across players and across time. Furthermore, we assume that the noise process $w_{i,t}$ has finite support for all players i and all times t.

The single period payoff to player i at time t is given as $\pi\left(x_{i,t}, a_{i,t}, \boldsymbol{x}_{-i,t}\right)$. Here $\boldsymbol{x}_{-i,t}$ is the state of all players except player i at time t. Note that the 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. Let us denote by $f_{-i,t}^{(m)}(y)$ the fraction of players (excluding player i) that have their state as y. We can define

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

where $\mathbf{1}_{\{x_{j,t}=y\}}$ is the indicator function that the state of player j at time t is y. Then, the payoff to player i depends on the state of other players via the distribution function $f_{-i,t}^{(m)}$. In a slight abuse of notation we will write the payoff to player i at time t as $\pi\big(x_{i,t}, a_{i,t}, f_{-i,t}^{(m)}\big)$. Such payoff functions are common in economics literature [13] and in wireless interference games [14]. From equation (1) and the definition of the payoff function, we note that the players are coupled via their payoff functions only.

We assume that the state update function h as well as the payoff function π are increasing and concave functions of the state x and action a. Furthermore, we assume that both state dynamics and payoff function have decreasing differences in state and action. We show that under these assumptions on the model primitives, the optimal oblivious policy is a decreasing function of state. This, along with the assumption that under zero action the system returns to zero state, leads to a stationary distribution which has all its moments finite and hence the light-tail condition holds. Having proved the light-tail, the remainder of the proof follows the developments given in [15]; this require some additional technical assumptions on the model. Thus, under these conditions, the oblivious equilibrium approximates the Markov perfect equilibrium.

REFERENCES

[1] 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, 2004.

[2] 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.

[3] R. Ericson and A. Pakes, "Markov-perfect industry dynamics: A framework for empirical work," *Review of Economic Studies*, vol. 62, no. 1, pp. 53–82, 1995.

[4] A. Pakes and P. McGuire, "Computing Markov-perfect Nash equilibria: Numerical implications of a dynamic differentiated product model," *RAND Journal of Economics*, vol. 25, no. 4, pp. 555–589, 1994.

[5] ——, "Stochastic algorithms, symmetric Markov perfect equilibrium, and the curse of dimensionality," *Econometrica*, vol. 69, no. 5, pp. 1261–1281, 2001.

[6] A. C. Antoulas and D. C. Sorensen, "Approximation of large-scale dynamical systems: An overview," *International Journal of Applied Mathematics and Computer Sciences*, vol. 11, no. 5, pp. 1093–1121, 2001.

[7] D. D. Siljak, *Decentralized Control of Complex Systems*. Academic Press, 1991.

[8] 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.

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

[10] G. Y. Weintraub, C. L. Benkard, and B. VanRoy, "Markov perfect industry dynamics with many firms," *Econometrica*, vol. 76, no. 6, 2008.

[11] 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.

[12] S. Adlakha, R. Johari, G. Weintraub, and A. Goldsmith, "Oblivious equilibrium for stochastic games with concave utility," *Proceedings of the Allerton Conference on Communication, Control and Computing*, 2007.

[13] G. Y. Weintraub, C. L. Benkard, and B. VanRoy, "Markov perfect industry dynamics with many firms," *Accepted to Econometrica*, 2007.

[14] J. Huang, R. Berry, and M. Honig, "Distributed interference compensation for wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 5, pp. 1074–1084, 2006.

[15] S. Adlakha, R. Johari, G. Weintraub, and A. Goldsmith, "Oblivious equilibrium for large-scale stochastic games with unbounded costs," *Proceedings of the IEEE Conference on Decision and Control*, 2008.