Oblivious Equilibrium for General Stochastic Games with Many Players

Vineet Abhishek\textsuperscript{1}, Sachin Adlakha\textsuperscript{2}, Ramesh Johari\textsuperscript{3} and Gabriel Weintraub\textsuperscript{4}

\textsuperscript{1) Department of Electrical Engineering, Stanford University.}
\textsuperscript{2) Department of Electrical Engineering, Stanford University.}
\textsuperscript{3) Department of Management Science and Engineering, Stanford University.}
\textsuperscript{4) Columbia Business School, Columbia University.}
Introduction

• We study a solution of large stochastic games.

• A common equilibrium notion for stochastic games is *Markov perfect equilibrium (MPE)*.

• In MPE, strategies of players depend on the current state of *all* players.

• MPE is used extensively in economics literature to study models of industry dynamics [Ericson and Pakes].
Related Work

• MPE is typically obtained numerically using dynamic programming.

• MPE computation is intractable for large problems [Pakes and McGuire].

• *Oblivious Equilibrium (OE)* was proposed for approximating MPE [Weintraub et al].

• In OE, a firm’s strategy depends only on its state and the long run average behavior of the industry.

• Oblivious equilibrium computation is significantly simpler than MPE.
Contributions

- The concept of OE was developed for an industry specific model and dynamics.

- We generalize the concept of OE solution for general class of stochastic games.
  - We have a more general action dependent payoff function.
  - We allow the possibility of heterogeneity in state evolution and payoff function.

- We isolate a set of parsimonious assumptions under which OE is a good approximation to MPE.
Model

- Consider an $m$ player stochastic game evolving over infinite horizon.

- Let $x_{i,t}$ be the state and $a_{i,t}$ be the action taken by player $i$ at time $t$.

- State evolution of a player $i$ is given by a conditional probability mass function as:

  $$x_{i,t+1} \sim h^{\theta_i}(x | x_{i,t}, a_{i,t})$$

  where $\theta_i$ is the type of the player.

- The single period payoff of a player $i$ given as

  $$\pi^{\theta_i} (x_{i,t}, a_{i,t}, f_{-i,t}(m), m),$$

  where, $f_{-i,t}(y)$ is the fraction of the players excluding player $i$ that have their state as $y$. 
Markov Perfect Equilibrium (MPE)

- Each player chooses an action $a_{i,t} = \mu_{m,\theta_i}(x_{i,t}, f_{-i,t})$ to maximize its expected present value.

- Value function for a player $i$ under a vector of control policy $\mu^m$ is given as

$$V_{\theta_i}^j(x, f, m| \mu^{m, \theta_i}, \mu^{m}_{-i}) \triangleq \mathbb{E} \left[ \sum_{\tau=t}^{\infty} \beta^{\tau-t} \pi_{\theta_i}^{j}(x_{i,\tau}, a_{i,\tau}, f_{-i,\tau}^{(m)}, m) \bigg| x_{i,t} = x, f_{-i,t}^{(m)} = f \right]$$

where $0 < \beta < 1$ is the discount factor.

- The vector of policies $\mu^m$ is an MPE if for all $j$, and $\mu' \in \mathcal{M}_{\theta_i}^j$, we have

$$\sup_{\mu'} V_{\theta_j}^j (x, f, m| \mu', \mu^m_{-j}) = V_{\theta_j}^j (x, f, m| \mu^{m, \theta_j}, \mu^{m}_{-j}) \ \forall x, f.$$  

- We focus on symmetric MPE, i.e., all players with same type $\theta$ use the same policy $\mu^{m, \theta}$. 
Oblivious Equilibrium (OE)

- In OE, a player $i$ chooses a policy $\tilde{\mu}^{m,\theta_i}$ that depends only on its current state and the average aggregate state of its competitors.

- Let $\tilde{f}^{(m)}$ denote the long-run average aggregate state of competitors and is defined as

$$\tilde{f}^{(m)}(y) \triangleq \mathbb{E} \left( f^{(m)}_{-i}(y) \right) = \frac{1}{m - 1} \sum_{j \neq i} q^{\mu^{m,\theta_j}}(y).$$

- We assume that the initial state of player $i$ is sampled from the stationary distribution $q^{\tilde{\mu}^{m,\theta_i}}$. 
Oblivious Equilibrium

- **Oblivious value function** for player $i$ is defined as

$$
\tilde{V}^\theta_i(x, m|\tilde{\mu}^{m,\theta_i}, \tilde{\mu}^{-i}_m) \triangleq E \left[ \sum_{\tau=t}^{\infty} \beta^{\tau-t} \pi^\theta_i (x_{i,\tau}, a_{i,\tau}, \tilde{f}^{(m)}_{-i}, m) \mid x_{i,t} = x; \tilde{\mu}^{m,\theta_i} \right].
$$

- The vector of policies $\tilde{\mu}^m$ represents an **oblivious equilibrium** if for all $j$, and $\mu' \in \tilde{M}^\theta_j$, we have

$$
\sup_{\mu'} \tilde{V}^\theta_j (x, m| \mu', \tilde{\mu}^-_m) = \tilde{V}^\theta_j (x, m| \tilde{\mu}^{m,\theta_j}, \tilde{\mu}^-_m), \ \forall x.
$$

Here $\tilde{M}^\theta$ is the set of all oblivious policies available to a player of type $\theta$. 
Asymptotic Markov Property (AME)

- The AME property says that oblivious policy is approximately optimal when compared against Markov policy.

- A sequence of oblivious policies $\tilde{\mu}_m$ possesses the asymptotic Markov equilibrium (AME) property if for all $x$ and $i$, and $\mu' \in M^{\theta_i}$, we have

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

• [A1] The Markov chain associated with the state evolution of each player \( i \) (with type \( \theta_i \)) playing an oblivious policy \( \tilde{\mu}_{m,\theta_i} \) is positive recurrent, and reaches a stationary distribution \( q_{\tilde{\mu}_{m,\theta_i}} \).

• [A2] Markov perfect equilibrium and oblivious equilibrium exist for the stochastic game under consideration.

• [A3] We assume that the payoff function is uniformly bounded. That is

\[
\sup_{x,a,(f(m),m)} \pi^\theta(x, a, f(m), m) < \infty \quad \forall \theta.
\]

Here \( \pi^\theta(x, a, f(m), m) \) is the payoff function for all players \( j \) with type \( \theta_j = \theta \).

• [A4] We assume that the payoff \( \pi^\theta \) is Gateaux differentiable with respect to \( f(m)(y) \).
Assumptions - Light Tail

- We define \( g^\theta(y) \) as the maximum rate of change of \( \pi^\theta \) with respect to a small change in fraction of competitors at state \( y \). That is

\[
g^\theta(y) \triangleq \sup_{x,a,f(m),m} \left| \frac{\partial \pi^\theta(x,a,f(m),m)}{\partial f(m)(y)} \right|
\]

- [A5] We assume that \( g^\theta(y) \) is finite for all \( \theta \) and \( y \). Also, given \( \epsilon > 0 \), \( \forall \theta \), there exists a state value \( z^\theta \), such that

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

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

- Light tail implies that the probability of competitors at larger state \( y \) goes to zero uniformly over \( m \).

- It also captures the effect of competitors at a higher state on the single period payoff of a player.
Main Theorem

Under the assumptions [A1]-[A5], a sequence of oblivious equilibrium policies $\tilde{\mu}^m$ satisfies the AME property. That is, for all $i, x$, and $\mu' \in \mathcal{M}^{\theta_i}$, we have

$$\lim_{m \to \infty} \mathbb{E} \left[ \sup_{\mu'} V^{\theta_i} (x, f, m \mid \mu', \tilde{\mu}^m_{-i}) - V^{\theta_i} (x, f, m \mid \tilde{\mu}^{m, \theta_i}, \tilde{\mu}^m_{-i}) \right] = 0.$$
Conclusions

• We generalized the concept of OE for stochastic games.

• Under mild technical conditions, the AME property holds and OE approximates MPE.

• OE allows analysis of problems with high dimension where MPE computation is intractable.

• For games with finite state space light tail condition trivially holds.