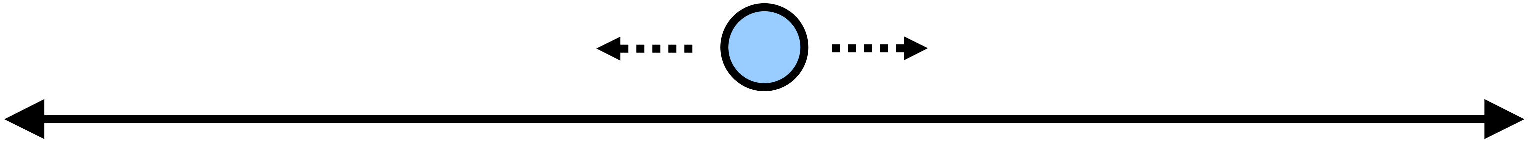
Mean-field equilibrium: An approximation approach for large dynamic games

Ramesh Johari

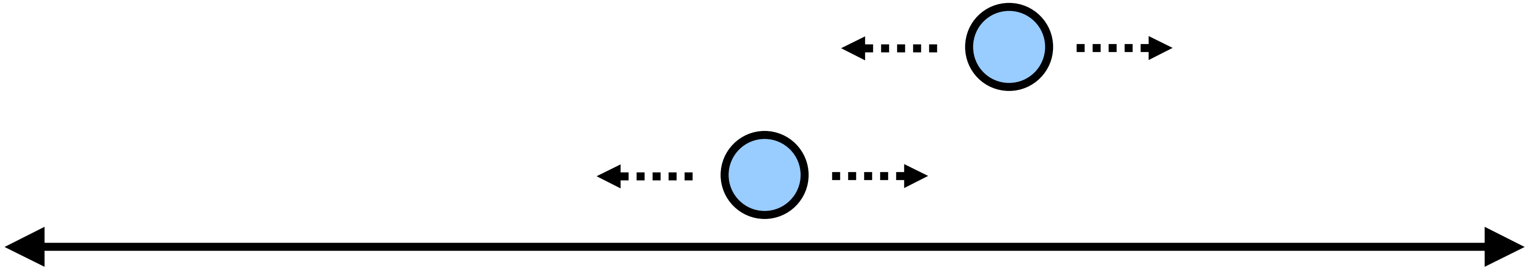
Stanford University

Sachin Adlakha, Gabriel Y. Weintraub, Andrea Goldsmith

Single agent dynamic control

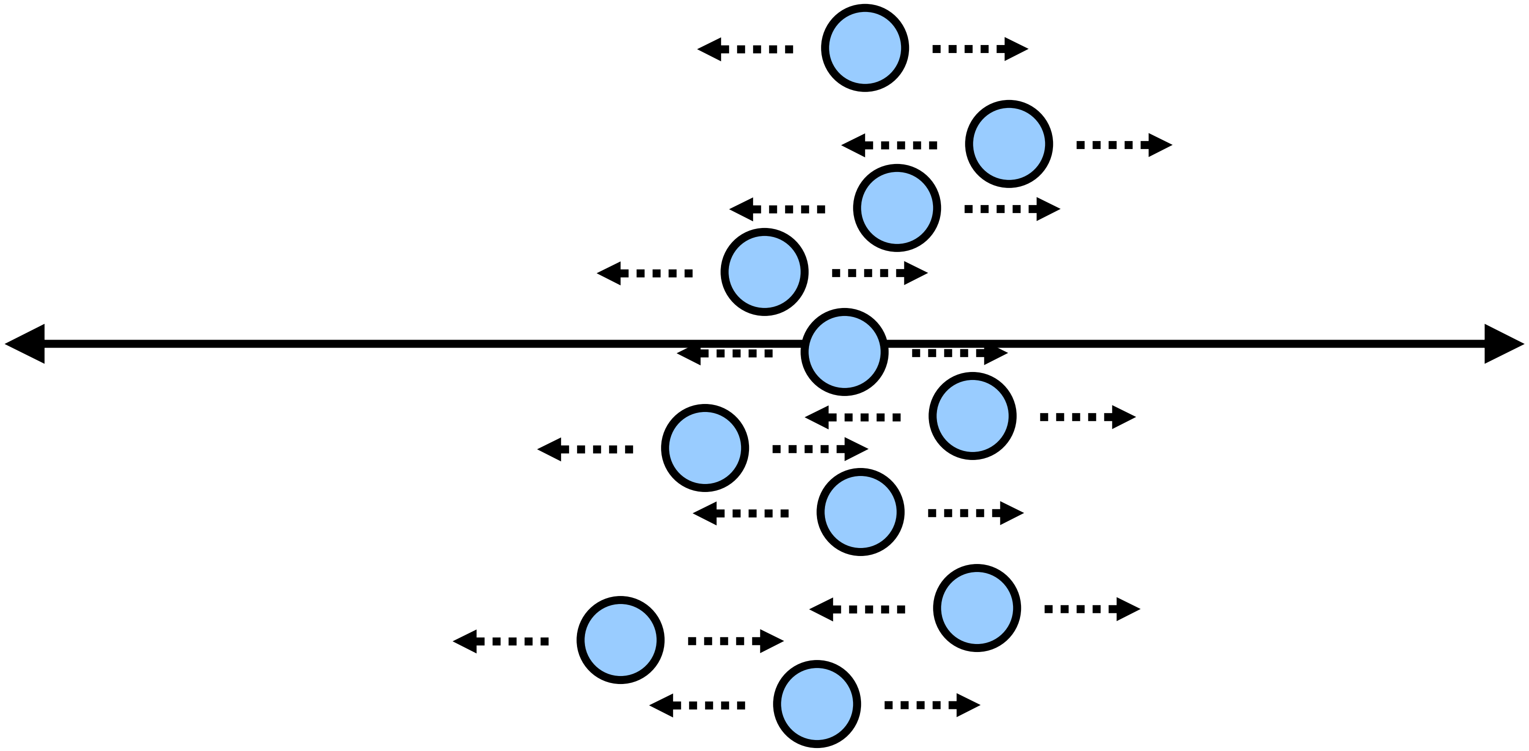


Two agents: a dynamic game



Suppose two devices are trying to *coordinate* on a common position.

Multiple agents: a dynamic game



Overview

- *Stochastic dynamic games:* Dynamic models of interacting self-interested agents
- *Large scale:* The number of agents is large

Traditional game theory is impractical in this regime:

- Equilibria in dynamic games make *very strong rationality assumptions*
- Equilibrium strategies grow in complexity with the number of agents (*"curse of dimensionality"*)

Mean field equilibrium (MFE)

- In large systems, it is plausible that agents can *ignore* the impact their actions have on behavior of others.
- We suppose agents treat empirical distribution of others as **fixed.**
- *Consistency check:* that distribution should arise from agents' optimal strategies.

This is the notion of *mean field equilibrium (MFE).*

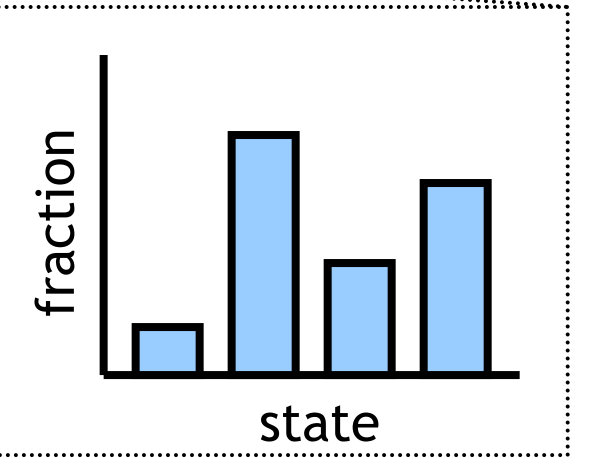
[Glynn et al.; Weintraub et al.; Huang et al.; Adlakha et al.; Yin et al.; Lasry and Lions; Bodoh-Creed; Duffie et al.]

Survey of results

1. **Approximation:**
   Under what conditions is MFE a good approximation to behavior in a finite system?
2. **Existence:**
   Under what conditions does a MFE exist?
3. **Convergence:**
   When do natural learning algorithms converge to MFE?

The model

- Discrete time, infinite horizon
- x_t : State of an agent at time t
- a_t : Action of an agent at time t
- $x_{t+1} \sim$ P($\cdot$ | x_t, a_t) : State transition kernel
- $\mathbf{f}_t$: Empirical distribution of others' states at time t (the *population state*)
- $\pi(x_t, \mathbf{f}_t) - c(a_t)$: Per period payoff
- β : Discount factor ($0 < \beta < 1$)



The model: example

Distributed coordination (Huang et al.):

Consider a collection of devices trying to coordinate to a common state.

Payoff: $\pi(x, \mathbf{f}) = -(x - \text{mean}(\mathbf{f}))^2$

Cost: $c(a) = a^2$

Dynamics: $x_{t+1} = A\, x_t + B\, a_t + w_t$, where w_t is i.i.d. noise

The model: generalizations

- Payoff dependent on actions of others
- Payoff, dynamics heterogeneous across agents
- Payoff nonseparable in state and action
- Dynamics coupled across agents
- Interaction with a finite subset
  of other players at each time step
- Nonstationarity

The model

An agent aims to maximize expected discounted payoff:

$$V(x|\mu', \mu) = \mathrm{E}\left[\sum_{t\geq 0} \beta^t (\pi(x_t, \mathbf{f}_t) - c(a_t)) \middle| x_0 = x, \mu', \mu\right]$$

μ' = this agent's strategy

μ = strategy followed by all others

Here μ' is a **cognizant strategy**:
it can depend on *both* x_t and $\mathbf{f}_t$

The mean field model

When the number of agents is large, suppose:

An agent reacts only to the long run average state distribution $\underline{f}$ of other players.

Mean field expected discounted payoff:

$$\underline{V}(x|\underline{\mu}) = \mathrm{E}\left[\sum_{t\geq 0}\beta^t(\pi(x_t, \underline{\mathbf{f}}) - c(a_t))\middle| x_0 = x, \underline{\mu}\right]$$

Here $\underline{\mu}$ is a **oblivious strategy**:
it depends *only* on x_t

Mean field equilibrium

A strategy $\underline{\mu}$ and a population state $\underline{\mathbf{f}}$ constitute a *mean field equilibrium (MFE)* if:

(1) $\underline{\mu}$ is an **optimal** oblivious strategy given $\underline{\mathbf{f}}$ and

(2) $\underline{\mathbf{f}}$ is a **steady state distribution** of $\underline{\mu}$

Distribution $\underline{\mathbf{f}}$ → Optimal strategy $\underline{\mu}$ for DP given $\underline{\mathbf{f}}$ → New steady state distribution $\Phi(\underline{\mathbf{f}})$

$\boxed{\text{A MFE population state } \underline{\mathbf{f}} \text{ is a } \textbf{fixed point} \text{ of } \Phi\text{: } \underline{\mathbf{f}} = \Phi(\underline{\mathbf{f}})}$

Our main insights

1. MFE is a good approximation under *continuity* and *compactness* conditions.
2. *Continuity* and *compactness* are (essentially) necessary conditions for existence of MFE in *concave games* and *supermodular games.*

   Thus approximation is a *corollary* of existence!
3. Technical challenges arise when the state space is unbounded or the dynamics are coupled, but similar results hold
4. In games with *complementarities*, simple learning dynamics converge to MFE

Ongoing applications

- Multi-armed bandit games
- Queueing games
- Dynamic resource allocation games

Asymptotic equilibrium

We first ask:

Is MFE a good approximation to equilibrium behavior of cognizant players?

A MFE ($\underline{\mu}$, $\underline{f}$) has the ***AE property*** if for all x, as number of players $\to \infty$,

$$\sup_{\mu'} V(x|\mu', \underline{\mu}) - V(x|\underline{\mu}, \underline{\mu}) \to 0,$$

where the sup is over all *cognizant strategies.*

[Weintraub et al.]

Asymptotic equilibrium

Suppose:

(1) State and action spaces are **compact**

(2) Payoff is **continuous** in f

Theorem:

The asymptotic equilibrium property holds for any MFE ($\underline{\mu}$, $\underline{f}$).

[Adlakha, Johari, Weintraub, Goldsmith]

Existence

The AE property can only hold if an MFE actually exists.

A central insight of this work is that the AE property is (essentially) a corollary of existence of MFE:

Existence also uses continuity and compactness.

Concave games

Assume compactness, continuity, and:

- Action space is convex
- π(x, f) concave in x
- $c(a)$ convex in a
- P($\cdot$ | x, a) has countable support, and is jointly stochastically concave in x, a

where at least one of the previous two is strict.

Concave games

Theorem:

There exists a MFE, and the AE property holds.

Note: via a slight variation, this result applies to finite state space, finite action space games.

[Adlakha, Johari, Weintraub, Goldsmith]