Mean-field Games for Fun and Profit

Santa Barbara Control Workshop

Decision, Dynamics and Control in Multi-Agent Systems

Sean P. Meyn

Joint work with: H. Yin, T. Yang, P. G. Mehta, and U. V. Shanbhag,

Coordinated Science Laboratory
and the Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign, USA

Thanks to NSF & AFOSR
Who Cares About Oscillators?
Background: Economics

Can you spot the lunacy?

A New Era For Control

We should remove derivative control from our engineering curriculum!

Fundamental Theorem of Calculus: If the airplane is flying at level height, then the ultimate contribution of the derivative is zero:

\[ 0 = y(T) - y(0) = \int_0^T \dot{y}(t) \, dt \]
Background: Economics
Can you spot the lunacy?

A New Era For Control
We should remove derivative control from our engineering curriculum!
A New Era For Control
We should remove derivative control from our engineering curriculum!

Fundamental Theorem of Calculus: If the airplane is flying at level height, then the ultimate contribution of the derivative is zero:

\[ 0 = y(T) - y(0) = \int_0^T \dot{y}(t) \, dt \]
Background: Economics
Can you spot the lunacy?

A New Era For Government
We should remove government spending from our economics curriculum!

Barro-Ricardo Equivalence Proposition: Government budget imbalances are irrelevant to resource allocation. Every dollar of taxes postponed today must be paid with interest tomorrow by the exact same group of taxpayers alive today.
Ericson and Pakes\footnote{Markov-perfect industry dynamics: A framework for empirical work, Rev. of Econ. Studies 1995}

This paper provides a model of firm and industry dynamics that allows for entry, exit and firm-specific uncertainty generating variability in the fortunes of firms. It focuses on the impact of uncertainty arising from investment in research and exploration-type processes. ... 

Coupled Markov models to address transients.

— Transients are \textit{everything} in both business and economics
Background: Economics

Saner voices

Ericson and Pakes\textsuperscript{[3]}

This paper provides a model of firm and industry dynamics that allows for entry, exit and firm-specific uncertainty generating variability in the fortunes of firms. It focuses on the impact of uncertainty arising from investment in research and exploration-type processes. ...

Coupled Markov models to address transients.

— Transients are \textit{everything} in both business and economics

\textit{Computation} of Nash equilibria for coupled MDP models?

\textsuperscript{[3]} Markov-perfect industry dynamics: A framework for empirical work, Rev. of Econ. Studies 1995
Background: Economics

Greater sanity

Weintraub, Benkard, and Van Roy\textsuperscript{[17]}

... oblivious equilibrium (OE) is an approximation in which each player makes decisions based on his own state and the “average” state of the other players. ...

Some aspects of dynamics and uncertainty are preserved.

Computation of Nash equilibria is possible.

\textsuperscript{[17]} Markov perfect industry dynamics with many firms, Econometrica 2008; \textsuperscript{[6]} Huang et al., TAC, 2007
Background: Synchrony in Nature

Synchrony is Good

Pacemaker cells

Laser light

Introduction
Background: Synchrony in Nature

Synchrony is Good

Pacemaker cells

Laser light

Synchrony is Not Good

Bridges

Brains
Question (Fundamental question in Neuroscience)

Why is synchrony (neural rhythms) useful?
Does it have a functional role?

Question (Fundamental question in Neuroscience)

Why is synchrony (neural rhythms) useful?
Does it have a functional role?

1 Synchronization

- Phase transition in controlled system (motivated by coupled oscillators)
- H. Yin, P. G. Mehta, S. P. Meyn and U. V. Shanbhag, “Synchronization of Coupled Oscillators is a Game,” TAC
Introduction

Question (Fundamental question in Neuroscience)

Why is synchrony (neural rhythms) useful?
Does it have a functional role?

1 Synchronization
   - Phase transition in controlled system (motivated by coupled oscillators)
   - H. Yin, P. G. Mehta, S. P. Meyn and U. V. Shanbhag, “Synchronization of Coupled Oscillators is a Game,” TAC

4 Learning
   - Synaptic plasticity via long term potentiation (Hebbian learning)
     “Neurons that fire together wire together”

Question (Fundamental question in Neuroscience)

**Why is synchrony (neural rhythms) useful?**
**Does it have a functional role?**

1. **Synchronization**
   - Phase transition in controlled system (motivated by coupled oscillators)
   - H. Yin, P. G. Mehta, S. P. Meyn and U. V. Shanbhag, “Synchronization of Coupled Oscillators is a Game,” TAC

2. **Learning**
   - Synaptic plasticity via long term potentiation (Hebbian learning)
     “Neurons that fire together wire together”

3. **Neuronal computations**
   - Bayesian inference
   - Neural circuits as particle filters (Lee & Mumford)

**Background: Kuramoto model**

\[
d\theta_i(t) = \left( \omega_i + \frac{\kappa}{N} \sum_{j=1}^{N} \sin(\theta_j(t) - \theta_i(t)) \right) dt + \sigma d\xi_i(t), \quad i = 1, \ldots, N
\]

Background: Kuramoto model

\[ d\theta_i(t) = \left( \omega_i + \frac{\kappa}{N} \sum_{j=1}^{N} \sin(\theta_j(t) - \theta_i(t)) \right) dt + \sigma \, d\xi_i(t), \quad i = 1, \ldots, N \]

- \( \omega_i \): taken from distribution \( g(\omega) \) over \([1 - \gamma, 1 + \gamma]\)
- \( \gamma \): measures the heterogeneity of the population

Background: Kuramoto model

\[d\theta_i(t) = \left(\omega_i + \frac{\kappa}{N} \sum_{j=1}^{N} \sin(\theta_j(t) - \theta_i(t))\right) dt + \sigma d\xi_i(t), \quad i = 1, \ldots, N\]

- \(\kappa\): measures the strength of coupling

Background: Kuramoto model

\[ d\theta_i(t) = \left( \omega_i + \frac{\kappa}{N} \sum_{j=1}^{N} \sin(\theta_j(t) - \theta_i(t)) \right) dt + \sigma d\xi_i(t), \quad i = 1, \ldots, N \]

\[ \omega_i: \text{taken from distribution } g(\omega) \text{ over } [1-\gamma, 1+\gamma] \]

\[ \gamma: \text{measures the heterogeneity of the population} \]

\[ \kappa: \text{measures the strength of coupling} \]

---

Oscillator Games

Mass

Influence

\[ \theta_i(t) \]

Oscillator Games
Oscillator Game

\(N\) oscillators with natural frequency \(\omega_i\), chosen from distribution \(g(\cdot)\)
Oscillator Game

\( N \) oscillators with natural frequency \( \omega_i \), chosen from distribution \( g(\cdot) \)

Dynamics of \( i^{th} \) oscillator,

\[
d\theta_i = (\omega_i + u_i(t)) \, dt + \sigma \, d\xi_i
\]
Oscillator Game

\( N \) oscillators with natural frequency \( \omega_i \), chosen from distribution \( g(\cdot) \)

Dynamics of \( i^{th} \) oscillator,
\[
d\theta_i = (\omega_i + u_i(t)) \, dt + \sigma \, d\xi_i
\]

Oscillator seeks control \( u_i(\cdot) \) to minimize,
\[
\eta_i(u_i; u_{-i}) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \mathbb{E}[ \left( c(\theta_i; \theta_{-i}) + \frac{1}{2} Ru_i^2 \right) ] \, ds
\]

- Cost of anarchy, \( c(\theta_i; \theta_{-i}) \)
- Cost of control, \( Ru_i^2 \)
Oscillator Game

\( N \) oscillators with natural frequency \( \omega_i \), chosen from distribution \( g(\cdot) \)

Dynamics of \( i^{th} \) oscillator, 
\[
d\theta_i = (\omega_i + u_i(t))\,dt + \sigma\,d\xi_i
\]

Oscillator seeks control \( u_i(\cdot) \) to minimize,
\[
\eta_i(u_i; u_{-i}) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \mathbb{E}[ c(\theta_i; \theta_{-i}) + \frac{1}{2} R u_i^2 ] \,ds
\]

Cost of anarchy,
\[
c(\theta_i; \theta_{-i}) = \frac{1}{N} \sum_{j \neq i} c^{\bullet}(\theta_i - \theta_j)
\]
Mean-field model

\[ d\theta_i = (\omega_i + u_i(t)) \, dt + \sigma \, d\xi_i \]

\[ \eta_i(u_i; u_{-i}) = \lim_{T \to \infty} \frac{1}{T} \int_0^T E[c(\theta_i; \theta_{-i}) + \frac{1}{2} R u_i^2] \, ds \]

\[ c(\theta_i; \theta_{-i}) = \frac{1}{N} \sum_{j \neq i} c^\bullet(\theta_i, \theta_j(t)) \xrightarrow{N \to \infty} \bar{c}(\theta_i, t) \]

---

Mean-field model

\[ d\theta_i = (\omega_i + u_i(t)) \, dt + \sigma \, d\xi_i \]

\[ \eta_i(u_i; u_{-i}) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \text{E}[c(\theta_i; \theta_{-i}) + \frac{1}{2} Ru_i^2] \, ds \]

Letting \( N \to \infty \) and assume \( c(\theta_i, \theta_{-i}) \to \bar{c}(\theta_i, t) \)

HJB: \[ \partial_t h + \omega \partial_{\theta} h = \frac{1}{2R} (\partial_{\theta} h)^2 - \bar{c}(\theta, t) + \eta^* - \frac{\sigma^2}{2} \partial_{\theta \theta}^2 h \Rightarrow h(\theta, t, \omega) \]

Mean-field model

\[ d\theta_i = (\omega_i + u_i(t)) dt + \sigma d\xi_i \]

\[ \eta_i(u_i; u_{-i}) = \lim_{T \to \infty} \frac{1}{T} \int_0^T E[c(\theta_i; \theta_{-i}) + \frac{1}{2} Ru_i^2] ds \]

Letting \( N \to \infty \) and assume \( c(\theta_i, \theta_{-i}) \to \bar{c}(\theta_i, t) \)

HJB: \[ \partial_t h + \omega \partial_\theta h = \frac{1}{2R} (\partial_\theta h)^2 - \bar{c}(\theta, t) + \eta^* - \frac{\sigma^2}{2} \partial_{\theta\theta}^2 h \quad \Rightarrow \quad h(\theta, t, \omega) \]

FPK: \[ \partial_t p + \omega \partial_\theta p = \frac{1}{R} \partial_\theta [p(\partial_\theta h)] + \frac{\sigma^2}{2} \partial_{\theta\theta}^2 p \quad \Rightarrow \quad p(\theta, t, \omega) \]

Mean-field model

\[ d\theta_i = (\omega_i + u_i(t)) dt + \sigma d\xi_i \]

\[ \eta_i(u_i; u_{-i}) = \lim_{T \to \infty} \frac{1}{T} \int_0^T E[c(\theta_i; \theta_{-i}) + \frac{1}{2} Ru_i^2] ds \]

Letting \( N \to \infty \) and assume \( c(\theta_i, \theta_{-i}) \to \bar{c}(\theta_i, t) \)

**HJB:** \[ \partial_t h + \omega \partial_{\theta} h = \frac{1}{2R} (\partial_{\theta} h)^2 - \bar{c}(\theta, t) + \eta^* - \frac{\sigma^2}{2} \partial_{\theta\theta}^2 h \Rightarrow h(\theta, t, \omega) \]

**FPK:** \[ \partial_t p + \omega \partial_{\theta} p = \frac{1}{R} \partial_{\theta} [p(\partial_{\theta} h)] + \frac{\sigma^2}{2} \partial_{\theta\theta}^2 p \Rightarrow p(\theta, t, \omega) \]

\[ \bar{c}(\vartheta, t) = \int_\Omega \int_0^{2\pi} c^\bullet(\vartheta, \theta) p(\theta, t, \omega) g(\omega) d\theta d\omega \]

Mean-field model

\[ d\theta_i = (\omega_i + u_i(t)) \, dt + \sigma \, d\xi_i \]

\[ \eta_i(u_i; u_{-i}) = \lim_{T \to \infty} \frac{1}{T} \int_0^T E[c(\theta_i; \theta_{-i}) + \frac{1}{2} Ru_i^2] \, ds \]

Letting \( N \to \infty \) and assume \( c(\theta_i, \theta_{-i}) \to \bar{c}(\theta_i, t) \)

HJB: \[ \partial_t h + \omega \partial_{\theta} h = \frac{1}{2R} (\partial_{\theta} h)^2 - \bar{c}(\theta, t) + \eta^* - \frac{\sigma^2}{2} \partial_{\theta \theta}^2 h \quad \Rightarrow \quad h(\theta, t, \omega) \]

FPK: \[ \partial_t p + \omega \partial_{\theta} p = \frac{1}{R} \partial_{\theta} [p(\partial_{\theta} h)] + \frac{\sigma^2}{2} \partial_{\theta \theta}^2 p \quad \Rightarrow \quad p(\theta, t, \omega) \]

\[ \bar{c}(\vartheta, t) = \int_\Omega \int_0^{2\pi} c^\bullet(\vartheta, \theta) \, p(\theta, t, \omega) \, g(\omega) \, d\theta \, d\omega \]

Solution of PDEs gives $\varepsilon$-Nash equilibrium
Solution of PDEs gives $\varepsilon$-Nash equilibrium

$\varepsilon$-Nash equilibrium

Solution to PDE $\Rightarrow$ Oblivious control for $i$th oscillator,

$u^o_i = -\frac{1}{R} \partial_\theta h(\theta(t), t, \omega) \bigg|_{\omega=\omega_i}$

Theorem: $\varepsilon$-Nash equilibrium property,

$\eta_i(u^o_i; u^o_{-i}) \leq \eta_i(u_i; u^o_{-i}) + O\left(\frac{1}{\sqrt{N}}\right), \quad i = 1, \ldots, N,$

for any adapted control $u_i$. 
Solution of PDEs gives $\varepsilon$-Nash equilibrium

$\varepsilon$-Nash equilibrium
Solution to PDE $\Rightarrow$ Oblivious control for $i$th oscillator,

$$u^o_i = -\frac{1}{R} \partial \theta h(\theta(t), t, \omega)|_{\omega=\omega_i}$$

Theorem: $\varepsilon$-Nash equilibrium property,

$$\eta_i(u^o_i; u^o_{-i}) \leq \eta_i(u_i; u^o_{-i}) + O\left(\frac{1}{\sqrt{N}}\right), \quad i = 1, \ldots, N,$$

for any adapted control $u_i$. 

Solution to PDE?
Incoherent Solution

\[ h(\theta, t, \omega) \equiv 0, \quad p(\theta, t, \omega) \equiv \frac{1}{2\pi} \]
Incoherent Solution

\[ h(\theta, t, \omega) \equiv 0, \quad p(\theta, t, \omega) \equiv \frac{1}{2\pi} \]

\[ h(\theta, t, \omega) = 0 \quad \Rightarrow \quad \partial_t h + \omega \partial_\theta h = \frac{1}{2R} (\partial_\theta h)^2 - \bar{c}(\theta, t) + \eta^* - \frac{\sigma^2}{2} \partial^2_{\theta\theta} h \]

\[ p(\theta, t, \omega) = \frac{1}{2\pi} \quad \Rightarrow \quad \partial_t p + \omega \partial_\theta p = \frac{1}{R} \partial_\theta [p(\partial_\theta h)] + \frac{\sigma^2}{2} \partial^2_{\theta\theta} p \]

\[ p(\theta, t, \omega) = \frac{1}{2\pi} \quad \Rightarrow \quad \bar{c}(\theta, t) = \int_{\Omega} \int_0^{2\pi} c^\bullet(\theta, \vartheta) p(\vartheta, t, \omega) g(\omega) d\vartheta d\omega \]

\[ = \int_{\Omega} \int_0^{2\pi} \frac{1}{2} \sin^2 \left( \frac{\theta - \vartheta}{2} \right) \frac{1}{2\pi} g(\omega) d\vartheta d\omega \]

\[ = \frac{1}{4} \]
Incoherent Solution

\[ h(\theta, t, \omega) \equiv 0, \quad p(\theta, t, \omega) \equiv \frac{1}{2\pi} \]

\[ h(\theta, t, \omega) = 0 \quad \Rightarrow \quad \partial_t h + \omega \partial_\theta h = \frac{1}{2R} (\partial_\theta h)^2 - \bar{c}(\theta, t) + \eta^* - \frac{\sigma^2}{2} \partial_{\theta\theta}^2 h \]

\[ p(\theta, t, \omega) = \frac{1}{2\pi} \quad \Rightarrow \quad \partial_t p + \omega \partial_\theta p = \frac{1}{R} \partial_\theta [p(\partial_\theta h)] + \frac{\sigma^2}{2} \partial_{\theta\theta}^2 p \]

\[ p(\theta, t, \omega) = \frac{1}{2\pi} \quad \Rightarrow \quad \bar{c}(\theta, t) = \int_\Omega \int_0^{2\pi} c^\bullet(\theta, \vartheta) p(\vartheta, t, \omega) g(\omega) \, d\vartheta \, d\omega \]

\[ = \int_\Omega \int_0^{2\pi} \frac{1}{2} \sin^2 \left( \frac{\theta - \vartheta}{2} \right) \frac{1}{2\pi} g(\omega) \, d\vartheta \, d\omega \]

\[ = \frac{1}{4} \]
Incoherent Solution

\[ h(\theta, t, \omega) \equiv 0, \quad p(\theta, t, \omega) \equiv \frac{1}{2\pi} \]

\[ h(\theta, t, \omega) = 0 \quad \Rightarrow \quad \partial_t h + \omega \partial_\theta h = \frac{1}{2R} (\partial_\theta h)^2 - \bar{c}(\theta, t) + \eta^* - \frac{\sigma^2}{2} \partial^2_{\theta\theta} h \]

\[ p(\theta, t, \omega) = \frac{1}{2\pi} \quad \Rightarrow \quad \partial_t p + \omega \partial_\theta p = \frac{1}{R} \partial_\theta [p(\partial_\theta h)] + \frac{\sigma^2}{2} \partial^2_{\theta\theta} p \]

\[ p(\theta, t, \omega) = \frac{1}{2\pi} \quad \Rightarrow \quad \bar{c}(\theta, t) = \int_{\Omega} \int_0^{2\pi} c^\bullet(\theta, \vartheta) p(\vartheta, t, \omega) g(\omega) \, d\vartheta \, d\omega \]

\[ = \int_{\Omega} \int_0^{2\pi} \frac{1}{2} \sin^2 \left( \frac{\theta - \vartheta}{2} \right) \frac{1}{2\pi} g(\omega) \, d\vartheta \, d\omega \]

\[ = \frac{1}{4} \]
Examples of Solutions: Incoherence and Synchrony

Incoherence

Synchrony
Bifurcation

\[
R^{-1/2}
\]

\begin{align*}
\frac{d\theta_i}{dt} &= (\omega_i + u_i) dt + \sigma d\xi_i \\
\eta_i(u_i; u_{-i}) &= \lim_{T \to \infty} \frac{1}{T} \int_0^T E[c(\theta_i; \theta_{-i}) + \frac{1}{2} Ru_i^2] ds
\end{align*}

Yin et al., ACC 2010
Bifurcation

\[ d\theta_i = (\omega_i + u_i) \, dt + \sigma \, d\xi_i \]

\[ \eta_i(u_i; u_{-i}) = \lim_{T \to \infty} \frac{1}{T} \int_0^T E[c(\theta_i; \theta_{-i}) + \frac{1}{2} R u_i^2] \, ds \]

Yin et al., ACC 2010

\[ d\theta_i = \left( \omega_i + \frac{\kappa}{N} \sum_{j=1}^N \sin(\theta_j - \theta_i) \right) \, dt + \sigma \, d\xi_i \]

Comparison of controls

Control law
\[ u_i = \varphi(\theta, t, \omega_i) := -\frac{1}{R} \partial_\theta h(\theta, t, \omega) \bigg|_{\omega=\omega_i} \]

![Diagram showing population density and control laws](image-url)
Comparison of controls

Control law

\[ u_i = \varphi(\theta, t, \omega_i) := -\frac{1}{R} \partial_\theta h(\theta, t, \omega) \bigg|_{\omega = \omega_i} \]

Mass Influence

Population Density

Cost 

\[ c^*(\theta) = \frac{1}{2} \sin^2 \left(\frac{\theta}{2}\right) \]

Control laws

- \(\omega = 1\)
- Kuramoto
Learning to Control
Approximate Dynamic Programming

- Optimality equation

\[
\min_{u_i} \left\{ c(\theta; \theta_{-i}(t)) + \frac{1}{2} Ru_i^2 + \mathcal{D}_{u_i} h_i(\theta, t) \right\} = \eta_i^* \\
=: H_i(\theta, u_i; \theta_{-i}(t))
\]

---

Watkins & Dayan, Q-learning, 1992; Bertsekas & Tsitsiklis, NDP, 1996; Mehta & Meyn, CDC 2009
Approximate Dynamic Programming

- Optimality equation
  \[
  \min_{u_i} \left\{ \underbrace{c(\theta; \theta_{-i}(t)) + \frac{1}{2} R u_i^2 + D_{u_i} h_i(\theta, t)}_{=: H_i(\theta, u_i; \theta_{-i}(t))} \right\} = \eta_i^*
  \]

- Optimal control law
  \[
  u_i^* = -\frac{1}{R} \partial_{\theta} h_i(\theta, t)
  \]

---

Watkins & Dayan, Q-learning, 1992; Bertsekas & Tsitsiklis, NDP, 1996; Mehta & Meyn, CDC 2009
Approximate Dynamic Programming

- **Optimality equation**
  \[
  \min_{u_i} \left\{ c(\theta; \theta_{-i}(t)) + \frac{1}{2} Ru_i^2 + \mathcal{D}_{u_i} h_i(\theta, t) \right\} = \eta_i^* \\
  =: H_i(\theta; u_i; \theta_{-i}(t))
  \]

- **Optimal control law**
  \[
  u^*_i = -\frac{1}{R} \partial_{\theta} h_i(\theta, t)
  \]

- **Parameterization for approximation:**
  \[
  H_i^{(A_i; \phi_i)}(\theta, u_i; \theta_{-i}(t)) = c(\theta; \theta_{-i}(t)) + \frac{1}{2} Ru_i^2 + (\omega_i - 1 + u_i) A_i S^{(\phi_i)} + \frac{\sigma^2}{2} A_i C^{(\phi_i)}
  \]

  where
  \[
  S^{(\phi)}(\theta, \theta_{-i}) = \frac{1}{N} \sum_{j \neq i} \sin(\theta - \theta_j - \phi), \quad C^{(\phi)}(\theta, \theta_{-i}) = \frac{1}{N} \sum_{j \neq i} \cos(\theta - \theta_j - \phi)
  \]

---

Watkins & Dayan, Q-learning, 1992; Bertsekas & Tsitsiklis, NDP, 1996; Mehta & Meyn, CDC 2009
Approximate Dynamic Programming

- Optimality equation
  \[
  \min_{u_i} \left\{ c(\theta; \theta_{-i}(t)) + \frac{1}{2} Ru_i^2 + \mathcal{D}_{u_i} h_i(\theta, t) \right\} = \eta^*_i
  =: H_i(\theta, u_i; \theta_{-i}(t))
  \]

- Optimal control law
  \[
  u^*_i = -\frac{1}{R} \partial_{\theta} h_i(\theta, t)
  \]

- Parameterization for approximation:
  \[
  H_i^{(A_i, \phi_i)}(\theta, u_i; \theta_{-i}(t)) = c(\theta; \theta_{-i}(t)) + \frac{1}{2} Ru_i^2 + (\omega_i - 1 + u_i) A_i S^{(\phi_i)} + \frac{\sigma^2}{2} A_i C^{(\phi_i)}
  \]
  where
  \[
  S^{(\phi)}(\theta, \theta_{-i}) = \frac{1}{N} \sum_{j \neq i} \sin(\theta - \theta_j - \phi), \quad C^{(\phi)}(\theta, \theta_{-i}) = \frac{1}{N} \sum_{j \neq i} \cos(\theta - \theta_j - \phi)
  \]

- Approx. optimal control:
  \[
  u^{(A_i, \phi_i)}_i = \arg \min_{u_i} \left\{ H_i^{(A_i, \phi_i)}(\theta, u_i; \theta_{-i}(t)) \right\} = -\frac{A_i}{RN} \sum_{j \neq i} \sin(\theta - \theta_j(t) - \phi_i)
  \]

Watkins & Dayan, Q-learning, 1992; Bertsekas & Tsitsiklis, NDP, 1996; Mehta & Meyn, CDC 2009
Learning algorithm

- Bellman error:

  Pointwise: \( \mathcal{L}(A_i, \phi_i)(\theta, t) = \min_{u_i} \{ H_i^{(A_i, \phi_i)} \} - \eta_i^{(A_i^*, \phi_i^*)} \)

- Stochastic approximation based on ODE,

  \[
  \tilde{e}(A_i, \phi_i) = \sum_{k=1}^{2} |\langle \mathcal{L}(A_i, \phi_i), \tilde{\phi}_k(\theta) \rangle|^2 \\
  \frac{dA_i}{dt} = -\varepsilon \frac{d\tilde{e}(A_i, \phi_i)}{dA_i}, \quad \frac{d\phi_i}{dt} = -\varepsilon \frac{d\tilde{e}(A_i, \phi_i)}{d\phi_i}
  \]
Learning algorithm

- **Bellman error:**
  \[ \mathcal{L}(A_i, \phi_i)(\theta, t) = \min_{u_i} \{ H_i(A_i, \phi_i) \} - \eta_i^{(A^*_i, \phi^*_i)} \]

- **Stochastic approximation based on ODE,**
  \[ \bar{e}(A_i, \phi_i) = \sum_{k=1}^{2} |\langle \mathcal{L}(A_i, \phi_i), \bar{\phi}_k(\theta) \rangle|^2 \]

  \[
  \begin{align*}
  \frac{dA_i}{dt} &= -\varepsilon \frac{d\bar{e}(A_i, \phi_i)}{dA_i}, \\
  \frac{d\phi_i}{dt} &= -\varepsilon \frac{d\bar{e}(A_i, \phi_i)}{d\phi_i}
  \end{align*}
  \]

Yin et al., CDC 2010
Comparison of average cost

\[ d\theta_i = (\omega_i + u_i)\, dt + \sigma \, d\xi_i; \quad u_i = -\frac{A_i^*}{RN} \sum_{j \neq i} \sin(\theta_i - \theta_j(t) - \phi_i^*) \]

\[ \eta_i(u_i; u_{-i}) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \mathbb{E}[c(\theta_i; \theta_{-i}) + \frac{1}{2} Ru_i^2] \, ds \]
Particle Filter Games
Filtering problem

Signal, observation processes:

\[ d\theta_t = \omega \, dt + \sigma_B \, dB_t \mod 2\pi \]
\[ dZ_t = h(\theta_t) \, dt + \sigma_W \, dW_t \]

\[ h(\theta) = \frac{1 + \cos(\theta)}{2} \]
Filtering problem

Signal, observation processes:

\[ d\theta_t = \omega \, dt + \sigma_B \, dB_t \mod 2\pi \]
\[ dZ_t = h(\theta_t) \, dt + \sigma_W \, dW_t \]

Nonlinear Filtering

Objective: estimate the posterior distribution \( p^* \) of \( \theta_t \) given \( \mathcal{Z}^t \).
Filtering problem

Signal, observation processes:

\[ d\theta_t = \omega \, dt + \sigma_B \, dB_t \mod 2\pi \]
\[ dZ_t = h(\theta_t) \, dt + \sigma_W \, dW_t \]

Nonlinear Filtering

Objective: estimate the posterior distribution \( p^* \) of \( \theta_t \) given \( Z^t \).

Solution approaches:

- Linear system: Kalman filter (R. E. Kalman, 1960)
- Nonlinear system: Wonham filter (W. M. Wonham, 1965)
- Numerical Methods: Particle filter (N. J. Gordon et al., 1993)
Feedback Particle Filter

Signal, observation processes:

\[ d\theta_t = \omega \, dt + \sigma_B \, dB_t \mod 2\pi \]
\[ dZ_t = h(\theta_t) \, dt + \sigma_W \, dW_t \]

Feedback Particle Filter

Particles evolve as controlled SDEs with independent noise,

\[ d\theta_t^i = \omega \, dt + \sigma_B \, dB_t^i + dU_t^i \mod 2\pi, \quad i = 1, \ldots, N. \]
Feedback Particle Filter

Signal, observation processes:

\[
d\theta_t = \omega \, dt + \sigma_B \, dB_t \mod 2\pi \\
dZ_t = h(\theta_t) \, dt + \sigma_W \, dW_t 
\]

Feedback Particle Filter

Particles evolve as controlled SDEs with independent noise,

\[
d\theta^i_t = \omega \, dt + \sigma_B \, dB^i_t + dU^i_t \mod 2\pi, \quad i = 1, \ldots, N.
\]

Objective: Choose control \( U^i_t \) so that,

\[
P\{\theta^i_t \in \cdot \mid Z^t_t\} = p^* = P\{\theta_t \in \cdot \mid Z^t_0\}
\]
Feedback Particle Filter

Signal, observation processes:

\[ d\theta_t = \omega \, dt + \sigma_B \, dB_t \mod 2\pi \]
\[ dZ_t = h(\theta_t) \, dt + \sigma_W \, dW_t \]

Feedback Particle Filter

Particles evolve as controlled SDEs with independent noise,

\[ d\theta^i_t = \omega \, dt + \sigma_B \, dB^i_t + dU^i_t \mod 2\pi, \quad i = 1, \ldots, N. \]

Objective: Choose control \( U^i_t \) so that,

\[ \mathbb{P}\{\theta^i_t \in \cdot \mid Z^t_0\} = p^* = \mathbb{P}\{\theta_t \in \cdot \mid Z^t_0\} \]

\( \implies \) Empirical distribution of particles approximates \( p^* \).
Filtering for Oscillator

Signal, observation processes:

\[ d\theta_t = \omega \, dt + \sigma_B \, dB_t \mod 2\pi \]
\[ dZ_t = h(\theta_t) \, dt + \sigma_W \, dW_t \]

Particle evolution,

\[ d\theta^i_t = \omega \, dt + \sigma_B \, dB^i_t + v(\theta^i_t)[dZ_t - \frac{1}{2}(h(\theta^i_t) + \hat{h}) \, dt] \]
\[ + \frac{1}{2} \sigma_W^2 \nu \nu' \, dt \mod 2\pi, \quad i = 1, \ldots, N. \]

Observer gain \( v(\theta^i_t) \) is obtained via the solution of an E-L equation,

\[ -\frac{\partial}{\partial \theta} \left( \frac{1}{p(\theta, t)} \frac{\partial}{\partial \theta} \left\{ p(\theta, t)v(\theta, t) \right\} \right) = -\frac{\sin \theta}{\sigma^2_W} \]
Filtering for Oscillator

Fourier form of $p(\theta, t)$,

$$p(\theta, t) = \frac{1}{2\pi} + P_s(t)\sin \theta + P_c(t)\cos \theta$$

Approx. solution of E-L equation, using a perturbation method:

$$v(\theta, t) = \frac{1}{2\sigma_w^2} \left\{ -\sin \theta + \frac{\pi}{2} [P_c(t)\sin 2\theta - P_s(t)\cos 2\theta] \right\} ,$$

$$v'(\theta, t) = \frac{1}{2\sigma_w^2} \left\{ -\cos \theta + \pi [P_c(t)\cos 2\theta + P_s(t)\sin 2\theta] \right\}$$

where

$$P_c(t) \approx \bar{P}_c^{(N)}(t) = \frac{1}{\pi N} \sum_{j=1}^{N} \cos \theta_t^j, \quad P_s(t) \approx \bar{P}_s^{(N)}(t) = \frac{1}{\pi N} \sum_{j=1}^{N} \sin \theta_t^j.$$
Simulation Results

Signal, observation processes:

\[ \begin{align*} 
  d\theta_t &= 1 \, dt + 0.5 \, dB_t \mod 2\pi \\
  dZ_t &= h(\theta_t) \, dt + 0.4 \, dW_t 
\end{align*} \]

\( N = 100 \) particles,

\[ d\theta^i_t = 1 \, dt + 0.5 \, dB^i_t + U(\theta^i_t; \bar{P}^{(N)}_c(t), \bar{P}^{(N)}_s(t)) \mod 2\pi \]

\[ h(\theta) = \frac{1 + \cos(\theta)}{2} \]
Simulation Results

Signal, observation processes:

\[
\begin{align*}
\text{d} \theta_t &= 1 \text{d}t + 0.5 \text{d}B_t \mod 2\pi \\
\text{d}Z_t &= h(\theta_t) \text{d}t + 0.4 \text{d}W_t
\end{align*}
\]

\( N = 100 \) particles,

\[
\text{d} \theta^i_t = 1 \text{d}t + 0.5 \text{d}B^i_t + U(\theta^i_t; \bar{P}_c^{(N)}(t), \bar{P}_s^{(N)}(t)) \mod 2\pi
\]
Simulation Results

Signal, observation processes:

\[
d\theta_t = 1 \, dt + 0.5 \, dB_t \mod 2\pi \\
dZ_t = h(\theta_t) \, dt + 0.4 \, dW_t
\]

\[N = 100\] particles,

\[
d\theta_t^i = 1 \, dt + 0.5 \, dB_t^i + U(\theta_t^i; \bar{P}_c^{(N)}(t), \bar{P}_s^{(N)}(t)) \mod 2\pi
\]
Variance Reduction
Filtering for simple linear model.

Mean-square error:
\[
\frac{1}{T} \int_0^T \left( \frac{\Sigma_t^{(N)} - \Sigma_t}{\Sigma_t} \right)^2 \, dt
\]
Conclusions
Fun and Profit?

Without doubt, MFGs provide a great playground. Neuro-morphic possibilities? What is the value of Winfree / Kuramoto models? I don't know. Possibilities for learning? Obvious!

The Feedback Particle Filter is a great playground, and has enormous potential for approximate nonlinear filtering in practice. Perhaps this is where the profit lies?
Conclusions

Fun and Profit?

Without doubt, MFGs provide a great playground.
Conclusions

Fun and Profit?

- Without doubt, MFGs provide a great playground.
- Neuro-morphic possibilities?
Without doubt, MFGs provide a great playground.

Neuro-morphic possibilities?
What is the value of Winfree / Kuramoto models?
I don’t know.
Conclusions

Fun and Profit?

- Without doubt, MFGs provide a great playground.
- Neuro-morphic possibilities?
  What is the value of Winfree / Kuramoto models?
  *I don’t know.*
- Possibilities for learning?
Conclusions

Fun and Profit?

- Without doubt, MFGs provide a great playground.
- Neuro-morphic possibilities?
  What is the value of Winfree / Kuramoto models?
  *I don’t know.*
- Possibilities for learning? *Obvious!*

The Feedback Particle Filter is a great playground, and has enormous potential for approximate nonlinear filtering in practice. Perhaps this is where the profit lies?
Conclusions

Fun and Profit?

- Without doubt, MFGs provide a great playground.
- Neuro-morphic possibilities?
  What is the value of Winfree / Kuramoto models? 
  *I don’t know.*
- Possibilities for learning? *Obvious!*

The Feedback Particle Filter is a great playground, and has enormous potential for approximate nonlinear filtering in practice.
Conclusions

Fun and Profit?

- Without doubt, MFGs provide a great playground.
- Neuro-morphic possibilities?
  What is the value of Winfree / Kuramoto models?
  I don’t know.
- Possibilities for learning? Obvious!

The Feedback Particle Filter is a great playground, and has enormous potential for approximate nonlinear filtering in practice.

Perhaps this is where the profit lies?
Thank you!

Collaborators

Huibing Yin  Tao Yang  Prashant Mehta  Uday Shanbhag

- “Q-learning and Pontryagin’s Minimum Principle,” CDC 2009
- “Synchronization of coupled oscillators is a game,” IEEE TAC, ACC 2010
- “Learning in Mean-Field Oscillator Games,” CDC 2010
- “On the Efficiency of Equilibria in Mean-field Oscillator Games,” ACC 2011
- “A Mean-field Control-oriented Approach for Particle Filtering,” ACC2011
- “Feedback Particle Filter with Mean-field Coupling,” CDC2011
Bibliography


