Mean-field Games for Fun and Profit

Santa Barbara Control Workshop
Decision, Dynamics and Control in Multi-Agent Systems

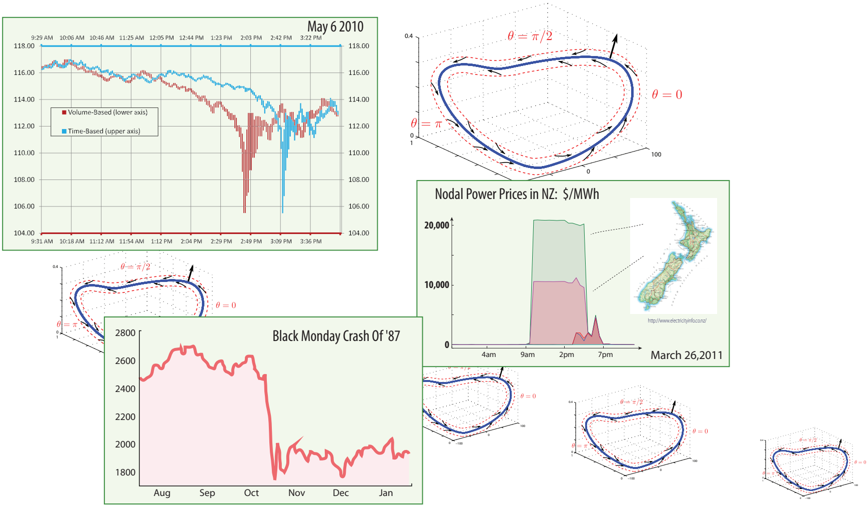
Sean P. Meyn

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

Outline



Who Cares About Oscillators?

Background: Economics

Can you spot the lunacy?

Background: Economics

Can you spot the lunacy?

A New Era For Control

We should remove derivative control from our engineering curriculum!

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 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.

Background: Economics

Saner voices

Ericson and Pakes [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 *everything* in both business and economics

[3] Markov-perfect industry dynamics: A framework for empirical work, Rev. of Econ. Studies 1995

Background: Economics

Saner voices

Ericson and Pakes [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 *everything* in both business and economics

Computation of Nash equilibria for coupled MDP models?

[3] Markov-perfect industry dynamics: A framework for empirical work, Rev. of Econ. Studies 1995

Background: Economics

Greater sanity

Weintraub, Benkard, and Van Roy [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.

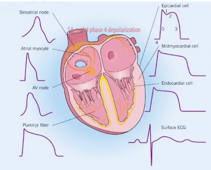
Computation of Nash equilibria is possible.

[17] Markov perfect industry dynamics with many firms, Econometrica 2008; [6] Huang et al., TAC, 2007

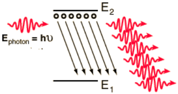
Background: Synchrony in Nature

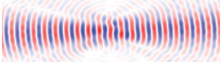
Synchrony is Good

Pacemaker cells



Laser light

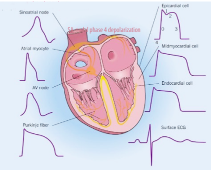




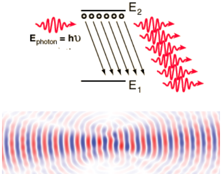
Background: Synchrony in Nature

Synchrony is Good

Pacemaker cells



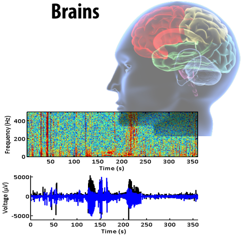
Laser light



Synchrony is Not Good

Bridges

Brains



Question (Fundamental question in Neuroscience)

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

Destexhe & Marder, Nature, 2004; Kopell et al., Neuroscience, 2009; Lee & Mumford, J. Opt. Soc, 2003

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

Destexhe & Marder, Nature, 2004; Kopell et al., Neuroscience, 2009; Lee & Mumford, J. Opt. Soc, 2003

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"
   - H. Yin, P. G. Mehta, S. P. Meyn and U. V. Shanbhag, " Learning in Mean-Field Oscillator Games," CDC 2010

Destexhe & Marder, Nature, 2004; Kopell et al., Neuroscience, 2009; Lee & Mumford, J. Opt. Soc, 2003

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

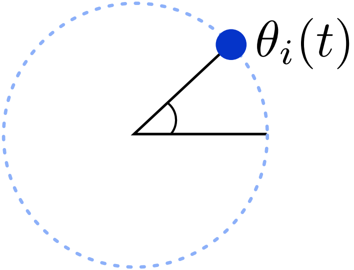
4. Learning
   - Synaptic plasticity via long term potentiation (Hebbian learning) "Neurons that fire together wire together"
   - H. Yin, P. G. Mehta, S. P. Meyn and U. V. Shanbhag, " Learning in Mean-Field Oscillator Games," CDC 2010

3. Neuronal computations
   - Bayesian inference
   - Neural circuits as particle filters (Lee & Mumford)
   - T. Yang, P. G. Mehta and S. P. Meyn, " A Control-oriented Approach for Particle Filtering," ACC&CDC 2011

Destexhe & Marder, Nature, 2004; Kopell et al., Neuroscience, 2009; Lee & Mumford, J. Opt. Soc, 2003

Background: Kuramoto model

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



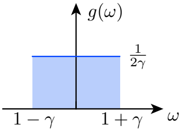
[9] Y. Kuramoto, 1975; [13] Strogatz et al., J. Stat. Phy., 1991

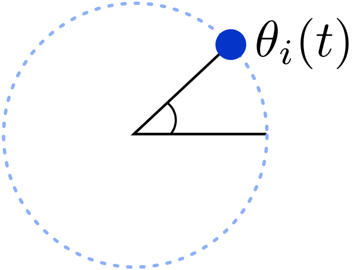
Background: Kuramoto model

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

- ω_i: taken from distribution $g(\omega)$ over $[1-\gamma, 1+\gamma]$

  γ: measures the heterogeneity of the population



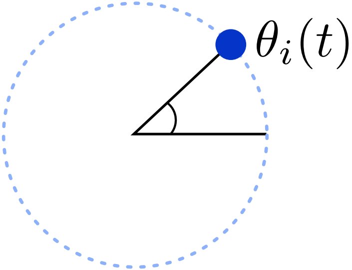


[9] Y. Kuramoto, 1975; [13] Strogatz et al., J. Stat. Phy., 1991

Background: Kuramoto model

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

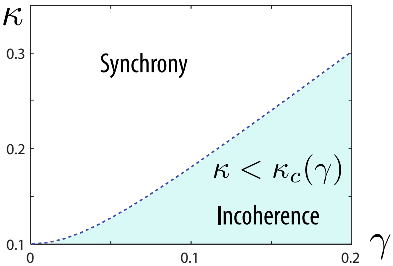
- $\boxed{\kappa\text{: measures the strength of coupling}}$



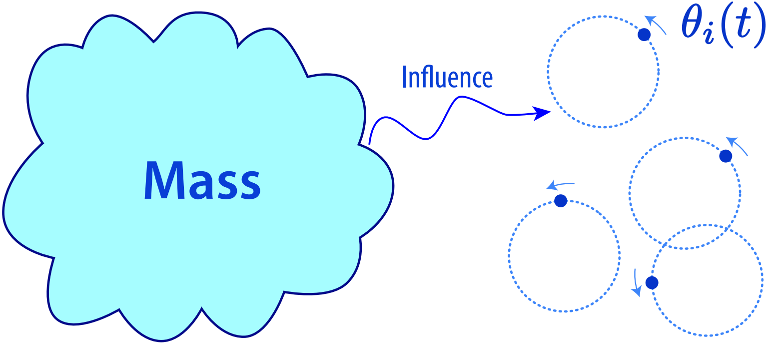
[9] Y. Kuramoto, 1975; [13] Strogatz et al., J. Stat. Phy., 1991

Background: Kuramoto model

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



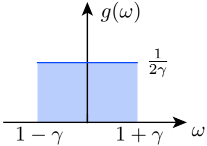
[9] Y. Kuramoto, 1975; [13] Strogatz et al., J. Stat. Phy., 1991



Oscillator Games

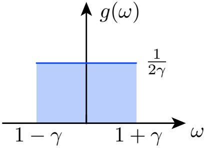
Oscillator Game

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



Oscillator Game

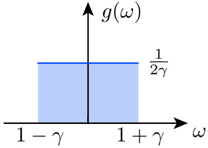
N oscillators with natural frequency ω_i, chosen from distribution $g(\cdot)$



Dynamics of i^{th} oscillator, $\quad \mathrm{d}\theta_i = (\omega_i + u_i(t))\,\mathrm{d}t + \sigma\,\mathrm{d}\xi_i$

Oscillator Game

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



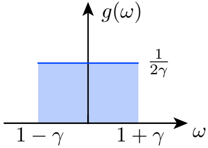
Dynamics of i^{th} oscillator, $\quad \mathrm{d}\theta_i = (\omega_i + u_i(t))\,\mathrm{d}t + \sigma\,\mathrm{d}\xi_i$

Oscillator seeks control $u_i(\cdot)$ to minimize,

$$\eta_i(u_i; u_{-i}) = \lim_{T\to\infty} \frac{1}{T}\int_0^T \mathsf{E}[\ \underbrace{c(\theta_i;\theta_{-i})}_{\text{cost of anarchy}} + \underbrace{\tfrac{1}{2}Ru_i^2}_{\text{cost of control}}\ ]\,\mathrm{d}s$$

Oscillator Game

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



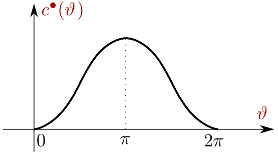
Dynamics of i^{th} oscillator, $\quad \mathrm{d}\theta_i = (\omega_i + u_i(t))\,\mathrm{d}t + \sigma\,\mathrm{d}\xi_i$

Oscillator seeks control $u_i(\cdot)$ to minimize,

$$\eta_i(u_i; u_{-i}) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathrm{E}[\ \underbrace{c(\theta_i;\theta_{-i})}_{\text{cost of anarchy}} + \underbrace{\tfrac{1}{2}Ru_i^2}_{\text{cost of control}}\ ]\,\mathrm{d}s$$

Cost of anarchy,

$$c(\theta_i;\theta_{-i}) = \frac{1}{N}\sum_{j\neq i} c^{\bullet}(\theta_i - \theta_j)$$



Mean-field model

$$\mathrm{d}\theta_i = (\omega_i + u_i(t))\,\mathrm{d}t + \sigma\,\mathrm{d}\xi_i$$

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

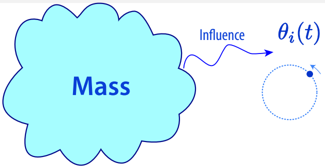
$$c(\theta_i;\theta_{-i}) = \frac{1}{N}\sum_{j\neq i} c^\bullet(\theta_i, \theta_j(t)) \xrightarrow{N\to\infty} \bar{c}(\theta_i, t)$$

[7] Huang et al., TAC, 2007; [12] Lasry & Lions, Japan. J. Math, 2007; [16] Weintraub et al., NIPS, 2006;

Mean-field model

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

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



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

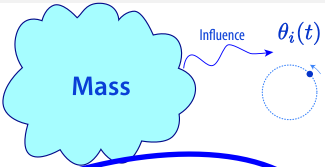
HJB: $\partial_t h + \omega\partial_\theta h = \dfrac{1}{2R}(\partial_\theta h)^2 - \bar{c}(\theta,t) + \eta^* - \dfrac{\sigma^2}{2}\partial^2_{\theta\theta} h \quad \Rightarrow \quad h(\theta,t,\omega)$

[7] Huang et al., TAC, 2007; [12] Lasry & Lions, Japan. J. Math, 2007; [16] Weintraub et al., NIPS, 2006;

Mean-field model

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

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



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

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

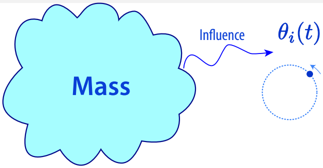
FPK: $\quad \partial_t p + \omega\partial_\theta p = \dfrac{1}{R}\partial_\theta[p(\partial_\theta h)] + \dfrac{\sigma^2}{2}\partial^2_{\theta\theta} p \quad \Rightarrow \quad p(\theta, t, \omega)$

[7] Huang et al., TAC, 2007; [12] Lasry & Lions, Japan. J. Math, 2007; [16] Weintraub et al., NIPS, 2006;

Mean-field model

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

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



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

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

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

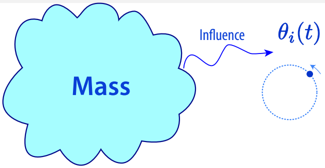
$$\bar{c}(\vartheta, t) = \int_\Omega \int_0^{2\pi} c^\bullet(\vartheta, \theta)\, p(\theta, t, \omega)\, g(\omega)\, d\theta\, d\omega$$

[7] Huang et al., TAC, 2007; [12] Lasry & Lions, Japan. J. Math, 2007; [16] Weintraub et al., NIPS, 2006;

Mean-field model

$$\mathrm{d}\theta_i = (\omega_i + u_i(t))\,\mathrm{d}t + \sigma\,\mathrm{d}\xi_i$$

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



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

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

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

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

[7] Huang et al., TAC, 2007; [12] Lasry & Lions, Japan. J. Math, 2007; [16] Weintraub et al., NIPS, 2006;

Solution of PDEs gives ε-Nash equilibrium

Solution of PDEs gives ε-Nash equilibrium

ε-Nash equilibrium

Solution to PDE $\Longrightarrow$ Oblivious control for ith oscillator,

$$\boxed{u_i^o = -\frac{1}{R}\partial_\theta h(\theta(t), t, \omega)\Big|_{\omega=\omega_i}}$$

Theorem: ε-Nash equilibrium property,

$$\eta_i(u_i^o; u_{-i}^o) \leq \eta_i(u_i; u_{-i}^o) + O(\frac{1}{\sqrt{N}}), \quad i = 1, \ldots, N,$$

for any adapted control u_i.

Solution of PDEs gives ε-Nash equilibrium

ε-Nash equilibrium

Solution to PDE $\Longrightarrow$ Oblivious control for ith oscillator,

$$\boxed{u_i^o = -\frac{1}{R}\partial_\theta h(\theta(t), t, \omega)\Big|_{\omega=\omega_i}}$$

Theorem: ε-Nash equilibrium property,

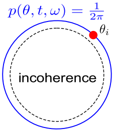
$$\eta_i(u_i^o; u_{-i}^o) \leq \eta_i(u_i; u_{-i}^o) + O(\frac{1}{\sqrt{N}}), \quad i = 1, \ldots, N,$$

for any adapted control u_i.

Solution to PDE?

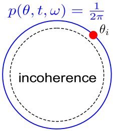
Incoherent Solution

$$h(\theta, t, \omega) \equiv 0\,, \qquad p(\theta, t, \omega) \equiv \frac{1}{2\pi}$$



Incoherent Solution

$$h(\theta, t, \omega) \equiv 0, \qquad p(\theta, t, \omega) \equiv \frac{1}{2\pi}$$



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

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

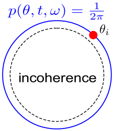
$$p(\theta, t, \omega) = \frac{1}{2\pi} \quad \Rightarrow \quad \bar{c}(\theta, t) = \int_\Omega \int_0^{2\pi} c^\bullet(\theta, \vartheta) p(\vartheta, t, \omega) g(\omega)\, \mathrm{d}\vartheta\, \mathrm{d}\omega$$

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

$$= \frac{1}{4}$$

Incoherent Solution

$$h(\theta,t,\omega) \equiv 0\,, \qquad p(\theta,t,\omega) \equiv \frac{1}{2\pi}$$



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

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

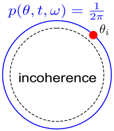
$$p(\theta,t,\omega) = \frac{1}{2\pi} \quad \Rightarrow \quad \bar{c}(\theta,t) = \int_\Omega \int_0^{2\pi} c^\bullet(\theta,\vartheta) p(\vartheta,t,\omega) g(\omega)\,\mathrm{d}\vartheta\,\mathrm{d}\omega$$

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

$$= \frac{1}{4}$$

Incoherent Solution

$$h(\theta,t,\omega) \equiv 0\,, \qquad p(\theta,t,\omega) \equiv \frac{1}{2\pi}$$



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

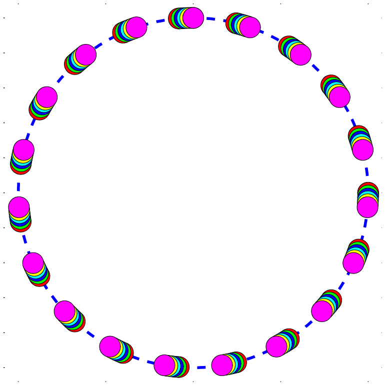
$$p(\theta,t,\omega)=\frac{1}{2\pi} \quad \Rightarrow \quad \partial_t p+\omega\partial_\theta p = \frac{1}{R}\partial_\theta[p(\partial_\theta h)] + \frac{\sigma^2}{2}\partial^2_{\theta\theta}p$$

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

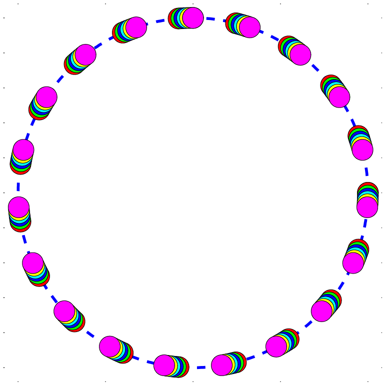
$$= \int_\Omega\int_0^{2\pi} \tfrac{1}{2}\sin^2\left(\frac{\theta-\vartheta}{2}\right)\frac{1}{2\pi}g(\omega)\,\mathrm{d}\vartheta\,\mathrm{d}\omega$$

$$= \frac{1}{4}$$

Examples of Solutions: Incoherence and Synchrony

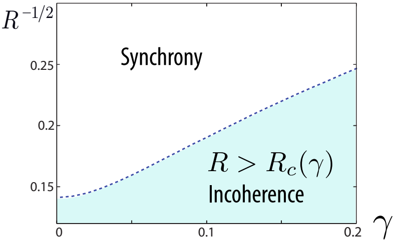


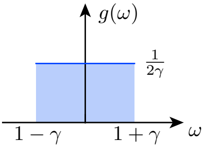
Incoherence



Synchrony

Bifurcation



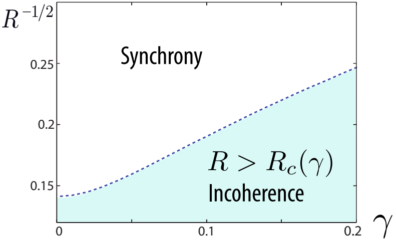


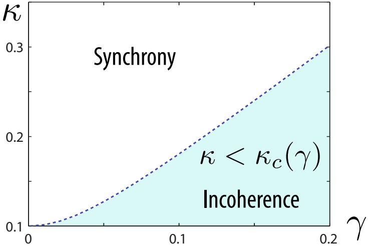
$$\mathrm{d}\theta_i = (\omega_i + u_i)\,\mathrm{d}t + \sigma\,\mathrm{d}\xi_i$$

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

Yin et al., ACC 2010

Bifurcation





$$\mathrm{d}\theta_i = (\omega_i + u_i)\,\mathrm{d}t + \sigma\,\mathrm{d}\xi_i$$

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

Yin et al., ACC 2010

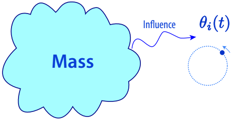
$$\mathrm{d}\theta_i = \left(\omega_i + \frac{\kappa}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i)\right)\mathrm{d}t + \sigma\,\mathrm{d}\xi_i$$

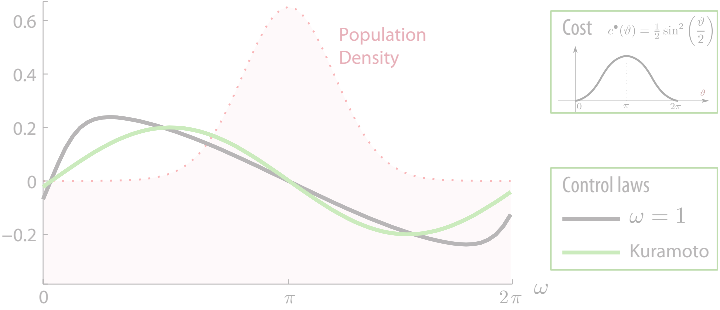
Strogatz et al., J. Stat. Phy., 1991

Comparison of controls

Control law

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

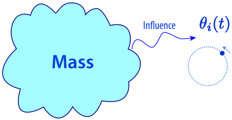


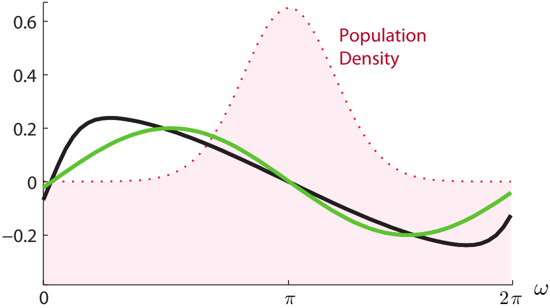


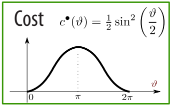
Comparison of controls

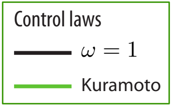
Control law

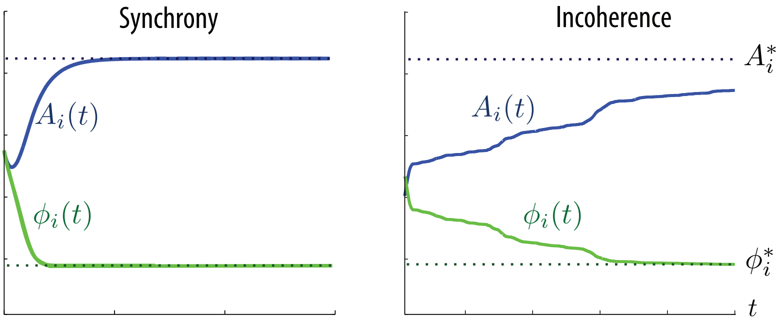
$$u_i = \varphi(\theta, t, \omega_i) := -\frac{1}{R}\partial_\theta h(\theta, t, \omega)\Big|_{\omega=\omega_i}$$











Mean-field Filter

Learning to Control

Approximate Dynamic Programming

- Optimality equation $\displaystyle \min_{u_i}\{\underbrace{c(\theta;\theta_{-i}(t))+\tfrac{1}{2}Ru_i^2+\mathscr{D}_{u_i}h_i(\theta,t)}_{=:\,H_i(\theta,u_i;\theta_{-i}(t))}\} = \eta_i^*$

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

Approximate Dynamic Programming

- Optimality equation $$\min_{u_i}\{\underbrace{c(\theta;\theta_{-i}(t))+\tfrac{1}{2}Ru_i^2+\mathscr{D}_{u_i}h_i(\theta,t)}_{=:\,H_i(\theta,u_i;\theta_{-i}(t))}\}=\eta_i^*$$

- 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 $\min_{u_i}\{\underbrace{c(\theta;\theta_{-i}(t))+\tfrac{1}{2}Ru_i^2+\mathscr{D}_{u_i}h_i(\theta,t)}_{=:\,H_i(\theta,u_i;\theta_{-i}(t))}\}=\eta_i^*$

- Optimal control law

$$u_i^* = -\frac{1}{R}\partial_\theta h_i(\theta,t)$$

- Parameterization for approximation:

$$H_i^{(A_i,\phi_i)}(\theta,u_i;\theta_{-i}(t)) = c(\theta;\theta_{-i}(t)) + \tfrac{1}{2}Ru_i^2 + (\omega_i - 1 + u_i)A_i S^{(\phi_i)} + \frac{\sigma^2}{2}A_i C^{(\phi_i)}$$

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 $\min_{u_i}\{\underbrace{c(\theta;\theta_{-i}(t))+\frac{1}{2}Ru_i^2+\mathscr{D}_{u_i}h_i(\theta,t)}_{=:\,H_i(\theta,u_i;\theta_{-i}(t))}\}=\eta_i^*$

- Optimal control law

$$u_i^* = -\frac{1}{R}\partial_\theta h_i(\theta,t)$$

- Parameterization for approximation:

$$H_i^{(A_i,\phi_i)}(\theta,u_i;\theta_{-i}(t)) = c(\theta;\theta_{-i}(t)) + \tfrac{1}{2}Ru_i^2 + (\omega_i - 1 + u_i)A_i S^{(\phi_i)} + \frac{\sigma^2}{2}A_i C^{(\phi_i)}$$

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:

$$u_i^{(A_i,\phi_i)} = \arg\min_{u_i}\{H_i^{(A_i,\phi_i)}(\theta,u_i;\theta_{-i}(t))\} = -\frac{A_i}{RN}\sum_{j\neq i}\sin(\theta-\theta_j(t)-\phi_i)$$

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

Learning algorithm

- Bellman error:

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

- Stochastic approximation based on ODE,

$$\tilde{e}(A_i,\phi_i) = \sum_{k=1}^{2} |\langle \mathscr{L}^{(A_i,\phi_i)}, \tilde{\varphi}_k(\theta)\rangle|^2$$

$$\frac{\mathrm{d}A_i}{\mathrm{d}t} = -\varepsilon \frac{\mathrm{d}\tilde{e}(A_i,\phi_i)}{\mathrm{d}A_i}, \quad \frac{\mathrm{d}\phi_i}{\mathrm{d}t} = -\varepsilon \frac{\mathrm{d}\tilde{e}(A_i,\phi_i)}{\mathrm{d}\phi_i}$$

Yin et al., CDC 2010

Learning algorithm

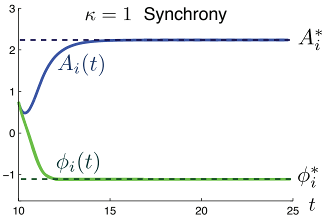
- Bellman error:

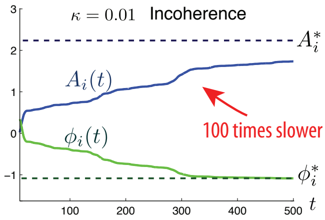
$$\text{Pointwise:} \quad \mathscr{L}^{(A_i,\phi_i)}(\theta,t) = \min_{u_i}\{H_i^{(A_i,\phi_i)}\} - \eta_i^{(A_i^*,\phi_i^*)}$$

- Stochastic approximation based on ODE,

$$\tilde{e}(A_i,\phi_i) = \sum_{k=1}^{2} |\langle \mathscr{L}^{(A_i,\phi_i)}, \tilde{\varphi}_k(\theta)\rangle|^2$$

$$\frac{\mathrm{d}A_i}{\mathrm{d}t} = -\varepsilon \frac{\mathrm{d}\tilde{e}(A_i,\phi_i)}{\mathrm{d}A_i}, \qquad \frac{\mathrm{d}\phi_i}{\mathrm{d}t} = -\varepsilon \frac{\mathrm{d}\tilde{e}(A_i,\phi_i)}{\mathrm{d}\phi_i}$$



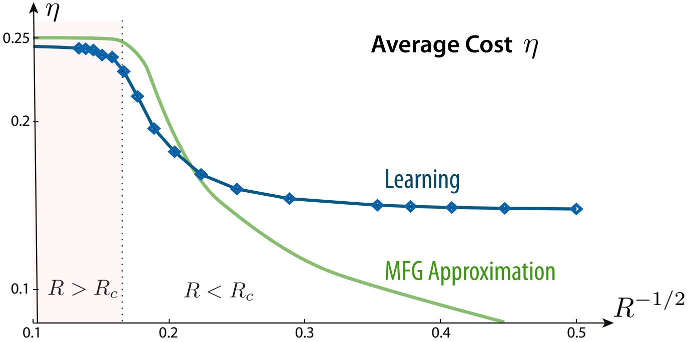


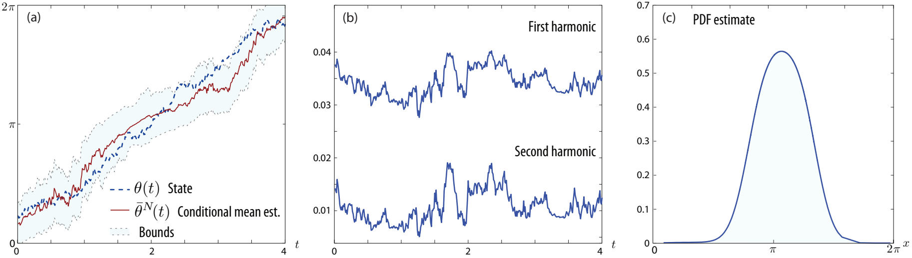
Yin et al., CDC 2010

Comparison of average cost

$$\mathrm{d}\theta_i = (\omega_i + u_i)\,\mathrm{d}t + \sigma\,\mathrm{d}\xi_i \qquad u_i = -\frac{A_i^*}{RN}\sum_{j\neq i}\sin(\theta_i - \theta_j(t) - \phi_i^*)$$

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



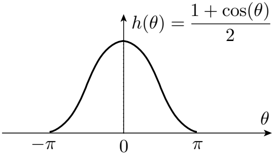


Particle Filter Games

Filtering problem

Signal, observation processes:

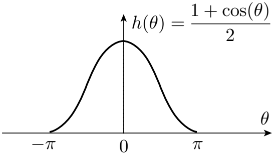
$$\mathrm{d}\theta_t = \omega\,\mathrm{d}t + \sigma_B\,\mathrm{d}B_t \quad \text{mod } 2\pi$$
$$\mathrm{d}Z_t = h(\theta_t)\,\mathrm{d}t + \sigma_W\,\mathrm{d}W_t$$



Filtering problem

Signal, observation processes:

$$\mathrm{d}\theta_t = \omega\,\mathrm{d}t + \sigma_B\,\mathrm{d}B_t \quad \mod 2\pi$$
$$\mathrm{d}Z_t = h(\theta_t)\,\mathrm{d}t + \sigma_W\,\mathrm{d}W_t$$



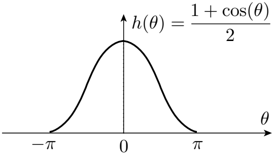
Nonlinear Filtering

Objective: estimate the posterior distribution p^* of θ_t given $\mathscr{Z}^t$.

Filtering problem

Signal, observation processes:

$$d\theta_t = \omega\, dt + \sigma_B\, dB_t \quad \text{mod } 2\pi$$
$$dZ_t = h(\theta_t)\, dt + \sigma_W\, dW_t$$



Nonlinear Filtering

Objective: estimate the posterior distribution p^* of θ_t given $\mathscr{Z}^t$.

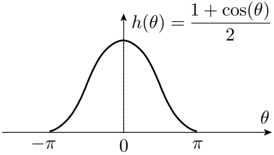
Solution approaches:

- 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:

$$\mathrm{d}\theta_t = \omega\,\mathrm{d}t + \sigma_B\,\mathrm{d}B_t \quad \text{mod } 2\pi$$
$$\mathrm{d}Z_t = h(\theta_t)\,\mathrm{d}t + \sigma_W\,\mathrm{d}W_t$$



Feedback Particle Filter

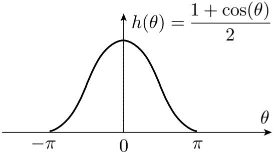
Particles evolve as controlled SDEs with independent noise,

$$\mathrm{d}\theta_t^i = \omega\,\mathrm{d}t + \sigma_B\,\mathrm{d}B_t^i + \mathrm{d}U_t^i \quad \text{mod } 2\pi, \qquad i = 1, \ldots, N.$$

Feedback Particle Filter

Signal, observation processes:

$$\begin{aligned} \mathrm{d}\theta_t &= \omega\,\mathrm{d}t + \sigma_B\,\mathrm{d}B_t \quad \text{mod } 2\pi \\ \mathrm{d}Z_t &= h(\theta_t)\,\mathrm{d}t + \sigma_W\,\mathrm{d}W_t \end{aligned}$$



Feedback Particle Filter

Particles evolve as controlled SDEs with independent noise,

$$\mathrm{d}\theta_t^i = \omega\,\mathrm{d}t + \sigma_B\,\mathrm{d}B_t^i + \mathrm{d}U_t^i \quad \text{mod } 2\pi, \qquad i = 1, \ldots, N.$$

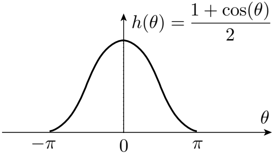
Objective: Choose control U_t^i so that,

$$\mathsf{P}\{\theta_t^i \in \cdot \mid Z_0^t\} = p^* = \mathsf{P}\{\theta_t \in \cdot \mid Z_0^t\}$$

Feedback Particle Filter

Signal, observation processes:

$$\mathrm{d}\theta_t = \omega\,\mathrm{d}t + \sigma_B\,\mathrm{d}B_t \quad \text{mod } 2\pi$$
$$\mathrm{d}Z_t = h(\theta_t)\,\mathrm{d}t + \sigma_W\,\mathrm{d}W_t$$



Feedback Particle Filter

Particles evolve as controlled SDEs with independent noise,

$$\mathrm{d}\theta_t^i = \omega\,\mathrm{d}t + \sigma_B\,\mathrm{d}B_t^i + \mathrm{d}U_t^i \quad \text{mod } 2\pi, \qquad i = 1, \ldots, N.$$

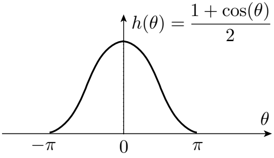
Objective: Choose control U_t^i so that,

$$\mathsf{P}\{\theta_t^i \in \cdot \mid Z_0^t\} = p^* = \mathsf{P}\{\theta_t \in \cdot \mid Z_0^t\}$$

$\Longrightarrow$ Empirical distribution of particles approximates p^*.

Filtering for Oscillator

Signal, observation processes:



$$
\begin{aligned}
\mathrm{d}\theta_t &= \omega\,\mathrm{d}t + \sigma_B\,\mathrm{d}B_t \quad \text{mod } 2\pi \\
\mathrm{d}Z_t &= h(\theta_t)\,\mathrm{d}t + \sigma_W\,\mathrm{d}W_t
\end{aligned}
$$

Particle evolution,

$$
\begin{aligned}
\mathrm{d}\theta_t^i = \omega\,\mathrm{d}t + \sigma_B\,\mathrm{d}B_t^i + v(\theta_t^i)[\mathrm{d}Z_t - \frac{1}{2}(h(\theta_t^i) + \hat{h})\,\mathrm{d}t] \\
+ \frac{1}{2}\sigma_W^2 v v'\,\mathrm{d}t \quad \text{mod } 2\pi, \qquad i = 1, \ldots, N.
\end{aligned}
$$

Observer gain $v(\theta_t^i)$ is obtained via the solution of an E-L equation,

$$
-\frac{\partial}{\partial\theta}\left(\frac{1}{p(\theta,t)}\frac{\partial}{\partial\theta}\{p(\theta,t)v(\theta,t)\}\right) = -\frac{\sin\theta}{\sigma_W^2}
$$

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:

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

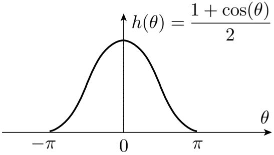
where

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

Simulation Results

Signal, observation processes:

$$\mathrm{d}\theta_t = 1\,\mathrm{d}t + 0.5\,\mathrm{d}B_t \quad \text{mod } 2\pi$$
$$\mathrm{d}Z_t = h(\theta_t)\,\mathrm{d}t + 0.4\,\mathrm{d}W_t$$

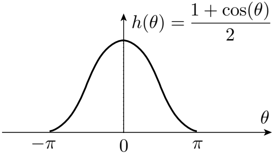


$N = 100$ particles,

$$\mathrm{d}\theta_t^i = 1\,\mathrm{d}t + 0.5\,\mathrm{d}B_t^i + U(\theta_t^i; \bar{P}_c^{(N)}(t), \bar{P}_s^{(N)}(t)) \quad \text{mod } 2\pi$$

Simulation Results

Signal, observation processes:

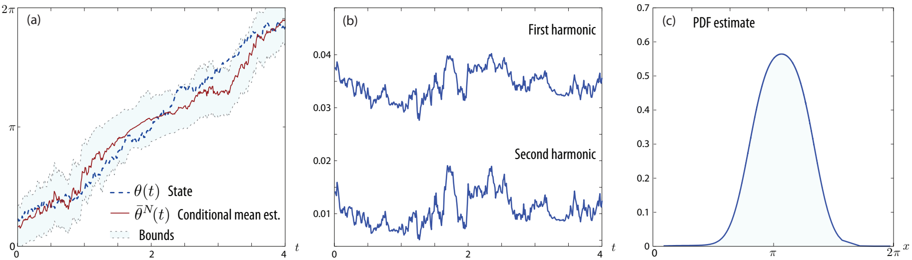


$$\mathrm{d}\theta_t = 1\,\mathrm{d}t + 0.5\,\mathrm{d}B_t \quad \text{mod } 2\pi$$

$$\mathrm{d}Z_t = h(\theta_t)\,\mathrm{d}t + 0.4\,\mathrm{d}W_t$$

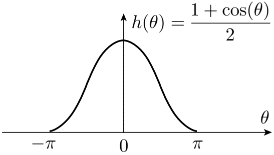
$N = 100$ particles,

$$\mathrm{d}\theta_t^i = 1\,\mathrm{d}t + 0.5\,\mathrm{d}B_t^i + U(\theta_t^i; \bar{P}_c^{(N)}(t), \bar{P}_s^{(N)}(t)) \quad \text{mod } 2\pi$$



Simulation Results

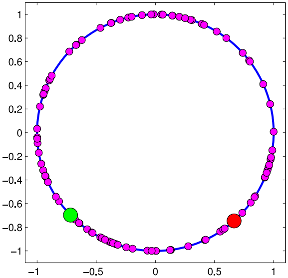
Signal, observation processes:



$$\mathrm{d}\theta_t = 1\,\mathrm{d}t + 0.5\,\mathrm{d}B_t \quad \text{mod } 2\pi$$
$$\mathrm{d}Z_t = h(\theta_t)\,\mathrm{d}t + 0.4\,\mathrm{d}W_t$$

$N = 100$ particles,

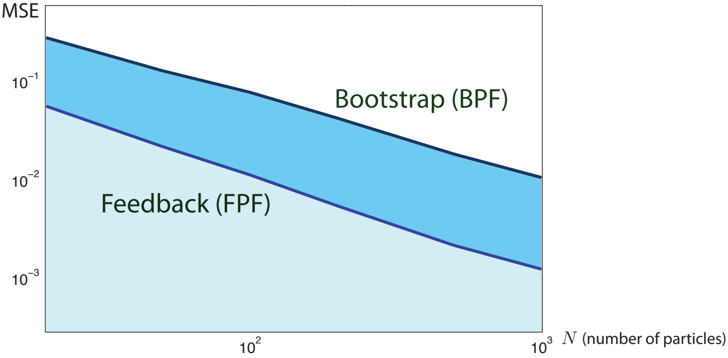
$$\mathrm{d}\theta^i_t = 1\,\mathrm{d}t + 0.5\,\mathrm{d}B^i_t + U(\theta^i_t; \bar{P}^{(N)}_c(t), \bar{P}^{(N)}_s(t)) \quad \text{mod } \ 2\pi$$



Variance Reduction

Filtering for simple linear model.

Mean-square error: $$\frac{1}{T}\int_0^T \left(\frac{\Sigma_t^{(N)} - \Sigma_t}{\Sigma_t}\right)^2 \mathrm{d}t$$



Conclusions

Fun and Profit?

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?

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.

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!

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.

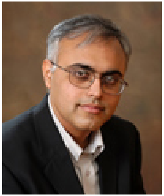
Perhaps this is where the profit lies?

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

Eric Brown, Jeff Moehlis, and Philip Holmes. On the phase reduction and response dynamics of neural oscillator populations. Neural Computation, 16(4):673–715, 2004.

A. Doucet, N. de Freitas, and N. Gordon. Sequential Monte-Carlo Methods in Practice. Springer-Verlag, April 2001.

R. Ericson and A. Pakes. Markov-perfect industry dynamics: A framework for empirical work. The Review of Economic Studies, 62(1):53–82, 1995.

N. J. Gordon, D. J. Salmond, and A. F. M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEE Proceedings F Radar and Signal Processing, 140(2):107–113, 1993.

J. Guckenheimer. Isochrons and phaseless sets. J. Math. Biol., 1:259–273, 1975.

M. Huang, P. E. Caines, and R. P. Malhame. Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε-Nash equilibria. 52(9):1560–1571, 2007.

Minyi Huang, Peter E. Caines, and Roland P. Malhame. Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε-nash equilibria. IEEE transactions on automatic control, 52(9):1560–1571, 2007.

R. E. Kalman. A new approach to linear filtering and prediction problems. Journal of Basic Engineering, 82(1):35–45, 1960.

Y. Kuramoto. International Symposium on Mathematical Problems in Theoretical Physics, volume 39 of Lecture Notes in Physics. Springer-Verlag, 1975.

- H. J. Kushner. On the differential equations satisfied by conditional probability densities of markov process. SIAM J. Control, 2:106–119, 1964.
- J. Lasry and P. Lions. Mean field games. Japanese Journal of Mathematics, 2(2):229–260, 2007.
- Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. Japan. J. Math., 2:229–260, 2007.
- S. H. Strogatz and R. E. Mirollo. Stability of incoherence in a population of coupled oscillators. Journal of Statistical Physics, 63:613–635, May 1991.
- Huibing Yin, Prashant G. Mehta, Sean P. Meyn, and Uday V. Shanbhag. Synchronization of coupled oscillators is a game. In Proc. of 2010 American Control Conference, pages 1783–1790, Baltimore, MD, 2010.
- Mehta P. G. Meyn S. P. Yin, H. and U. V. Shanbhag. Synchronization of coupled oscillators is a game. IEEE Trans. Automat. Control.
- G. Y. Weintraub, L. Benkard, and B. Van Roy. Oblivious equilibrium: A mean field approximation for large-scale dynamic games. In Advances in Neural Information Processing Systems, volume 18. MIT Press, 2006.
- G. Y. Weintraub, L. Benkard, and B. V. Roy. Markov perfect industry dynamics with many firms. Econometrica, 76(6):1375–1411, 2008.