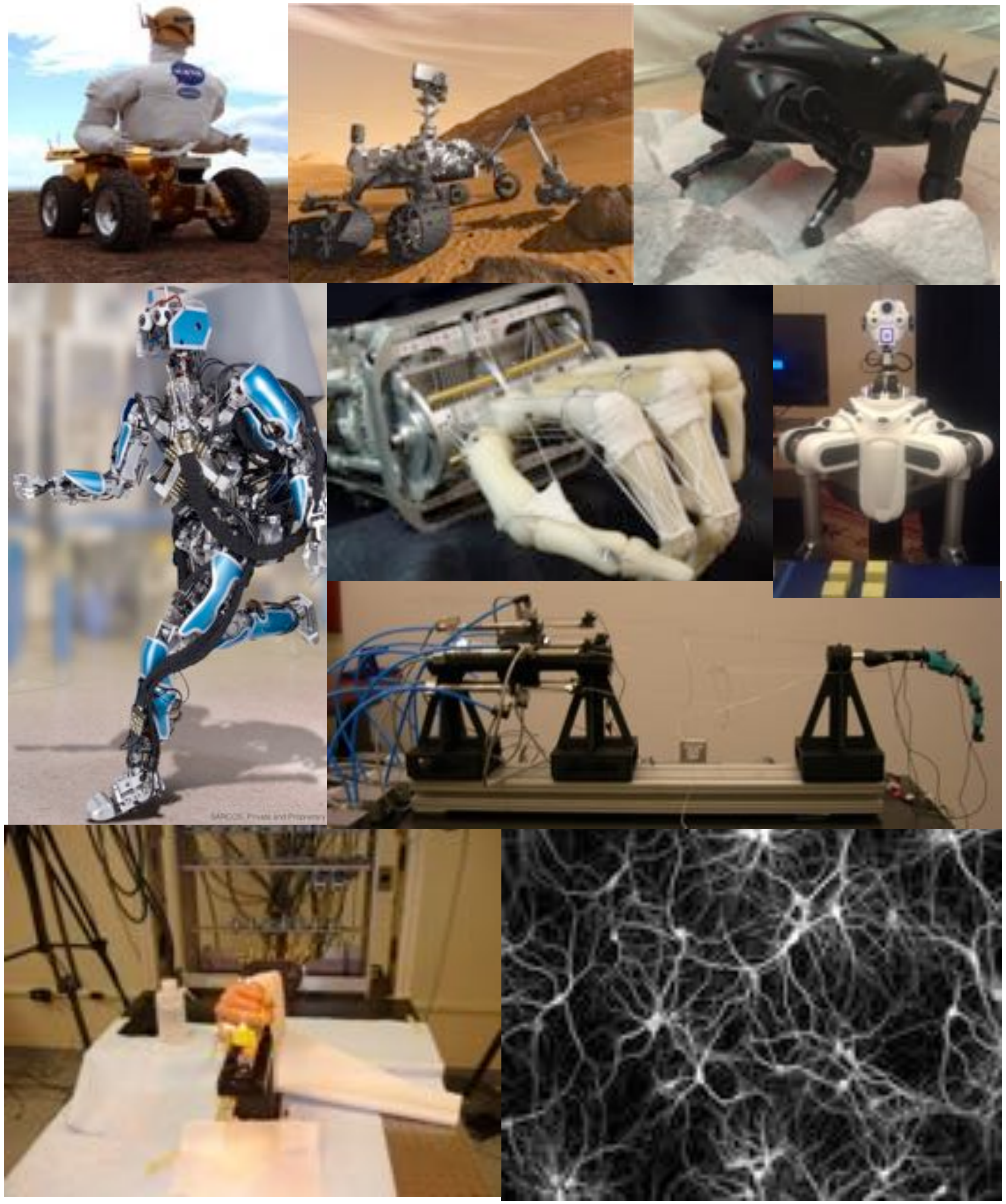
Real Time Stochastic Control and Decision Making: From theory to algorithms and applications

Evangelos A. Theodorou
Autonomous Control and Decision Systems Lab
Georgia Institute of Technology

Challenges in control



✦ Uncertainty

Stochastic uncertainty

Parametric uncertainty

Stochastic and parametric uncertainty

Total uncertainty

✦ Scalability

State space

Control parameterization

✦ Computational Efficiency

Real time

Few data/interactions

✦ Optimality

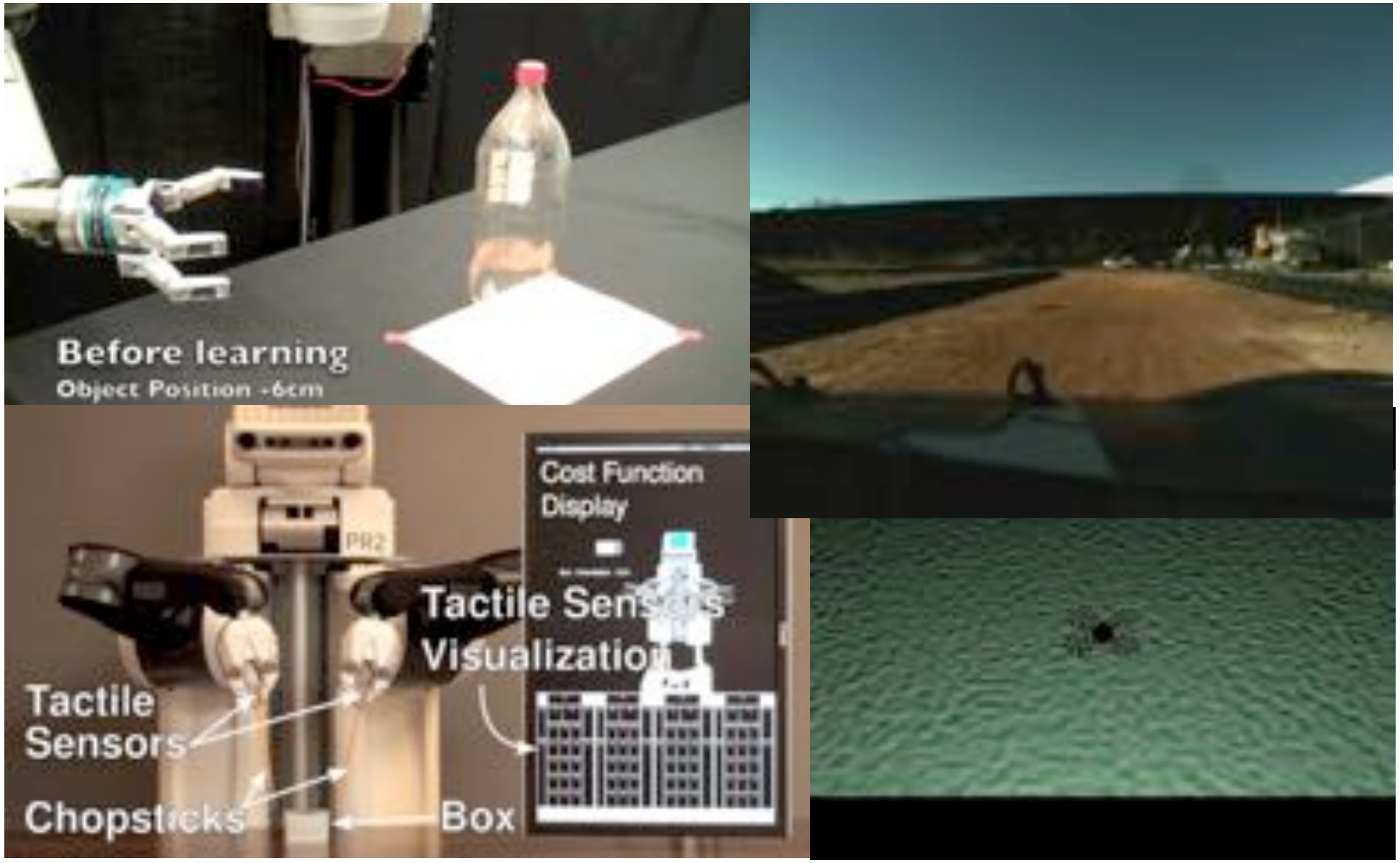
Global optimality

First principles

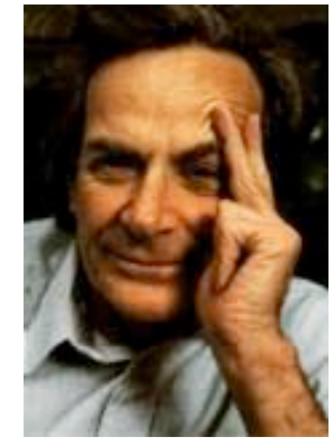
Stochastic Control Theory

Richard Bellman

Lev Pontryagin



Statistical Physics

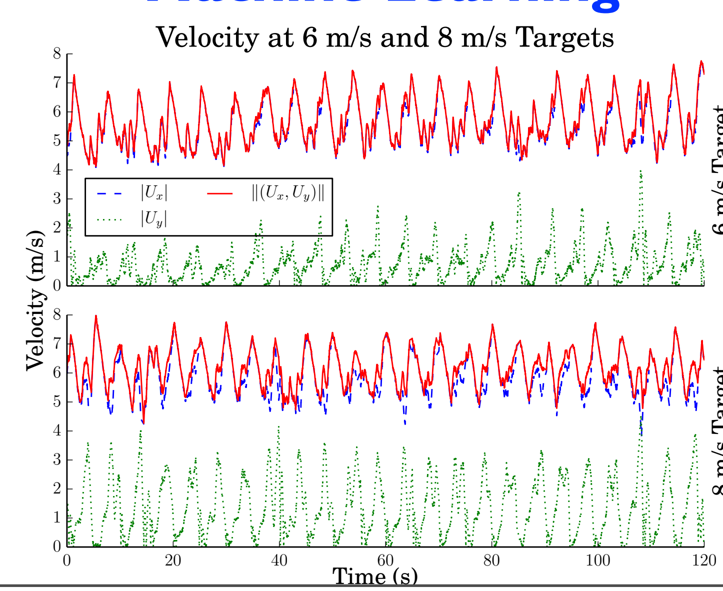


Richard Feynman

Mark Kac

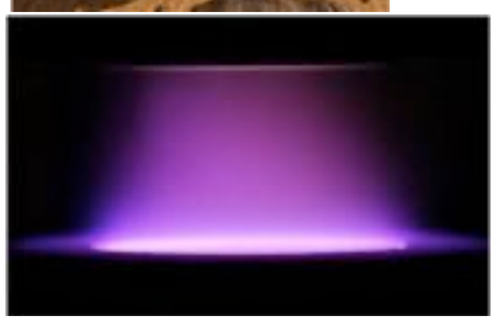
Parallel Computation

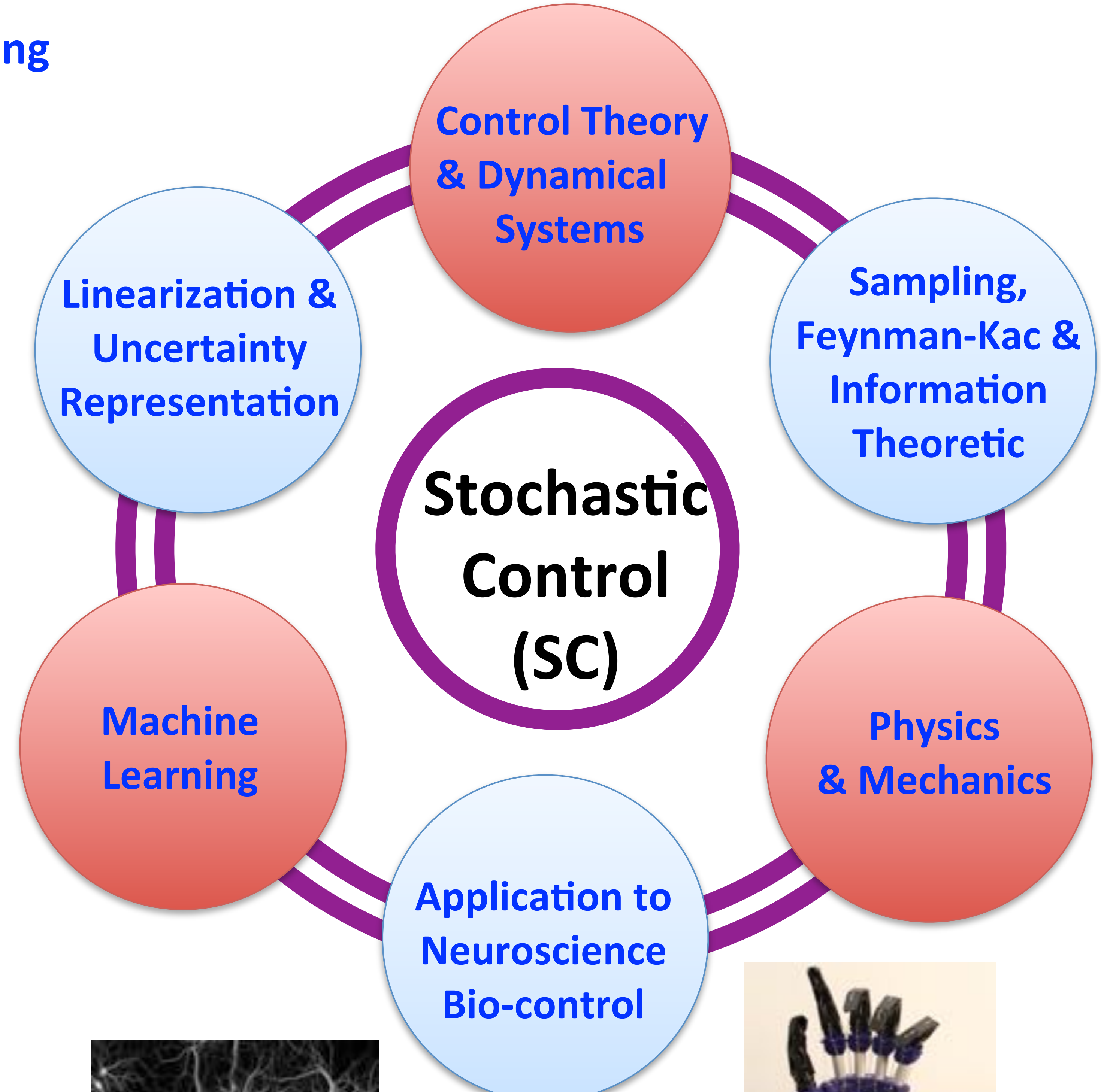
Machine Learning



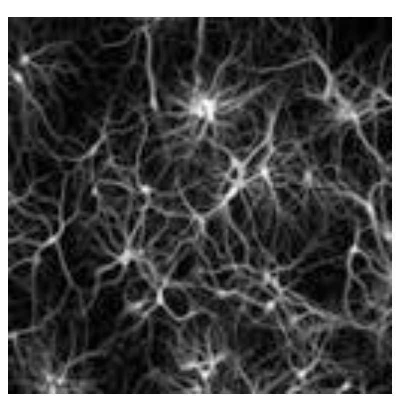
Autonomy

Aerospace Engineering





Robotics



Neuroscience & Bio-inspired Robotics

Theoretical Areas

Dynamic Programming

Problem Formulation:

Cost under minimization:

$$V(\mathbf{x}, t_i) = \min_{\mathbf{u}} \mathbb{E}_{\mathbb{Q}}\left[\phi(\mathbf{x}, t_N) + \int_{t_i}^{t_N} \mathcal{L}(\mathbf{x}, \mathbf{u}, t)\mathrm{d}t\right]$$

Running Cost:

$$\mathcal{L}(\mathbf{x}, \mathbf{u}, t) = q_0(\mathbf{x}, t) + q_1(\mathbf{x}, t)\mathbf{u} + \frac{1}{2}\mathbf{u}^T\mathbf{R}\mathbf{u}$$

Controlled Diffusion Dynamics:

$$\mathrm{d}\mathbf{x} = F(\mathbf{x}, \mathbf{u})\mathrm{d}t + \mathbf{B}(\mathbf{x})\mathrm{d}\omega$$

Noise

$$\mathbf{F}(\mathbf{x}, \mathbf{u}) = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u}$$

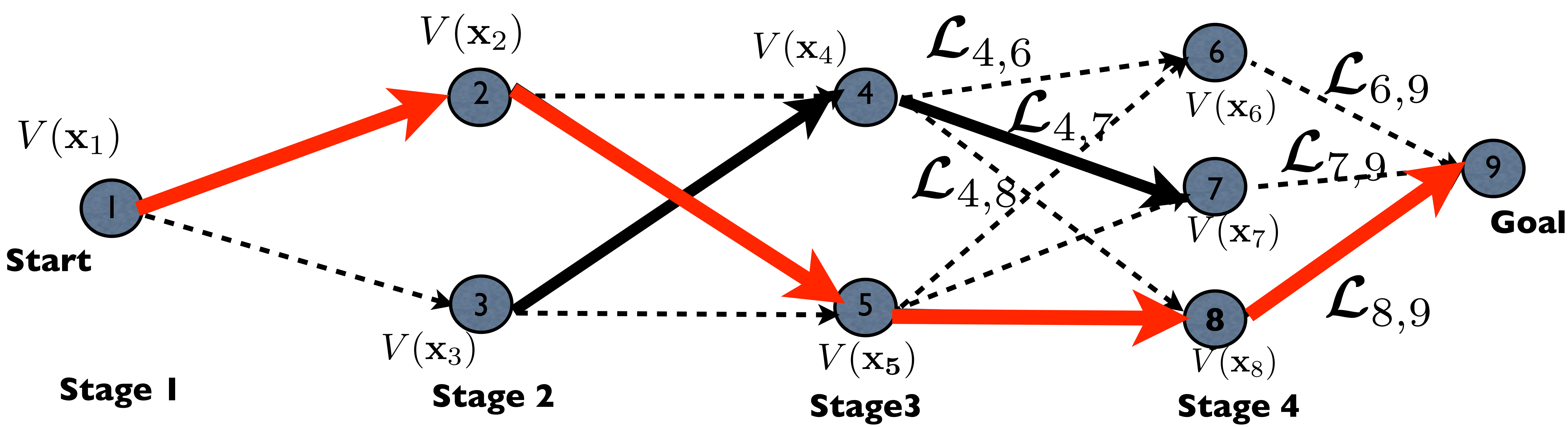
Control

$\mathbb{E}_{\mathbb{Q}}$ **: Expectation w.r.t controlled dynamics**

$\mathbb{E}_{\mathbb{P}}$ **: Expectation w.r.t uncontrolled dynamics**

Bellman Principle Discrete:

Cost to go(current state**) = min[Cost to reach(**next state**) + Cost to go(**next state**)]**



Dynamic Programming

Problem Formulation:

Cost under minimization:

$$V(\mathbf{x}, t_i) = \min_{\mathbf{u}} \mathbb{E}_{\mathbb{Q}} \left[\phi(\mathbf{x}, t_N) + \int_{t_i}^{t_N} \mathcal{L}(\mathbf{x}, \mathbf{u}, t) \mathrm{d}t \right]$$

Running Cost:

$$\mathcal{L}(\mathbf{x}, \mathbf{u}, t) = q_0(\mathbf{x}, t) + q_1(\mathbf{x}, t)\mathbf{u} + \frac{1}{2}\mathbf{u}^T \mathbf{R}\mathbf{u}$$

Controlled Diffusion Dynamics:

$$\mathrm{d}\mathbf{x} = F(\mathbf{x}, \mathbf{u})\mathrm{d}t + \mathbf{B}(\mathbf{x})\mathrm{d}\omega$$

Noise

$$\mathbf{F}(\mathbf{x}, \mathbf{u}) = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u}$$

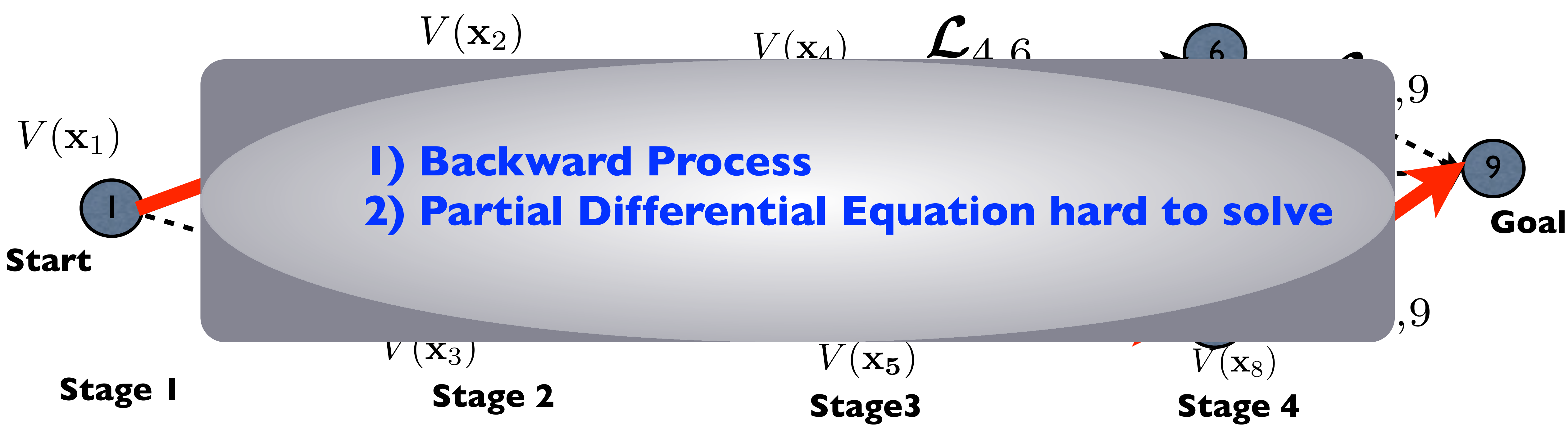
Control

$\mathbb{E}_{\mathbb{Q}}$ **: Expectation w.r.t controlled dynamics**

$\mathbb{E}_{\mathbb{P}}$ **: Expectation w.r.t uncontrolled dynamics**

Bellman Principle Discrete:

Cost to go(current state**) = min[Cost to reach(**next state**) + Cost to go(**next state**)]**



Making Control Linear

Hamilton Jacobi Bellman:

$$-\partial_t V = \tilde{q} + (\nabla_{\mathbf{x}} V)^T \tilde{\mathbf{f}} - \frac{1}{2}(\nabla_{\mathbf{x}} V)^T \mathbf{G}\mathbf{R}^{-1}\mathbf{G}^T(\nabla_{\mathbf{x}} V) + \frac{1}{2} tr\left((\nabla_{\mathbf{xx}} V)\mathbf{B}\mathbf{B}^T\right)$$

Desirability:

$$\Psi(\mathbf{x}, t) = \exp\left(-\frac{1}{\lambda}V(\mathbf{x}, t)\right)$$

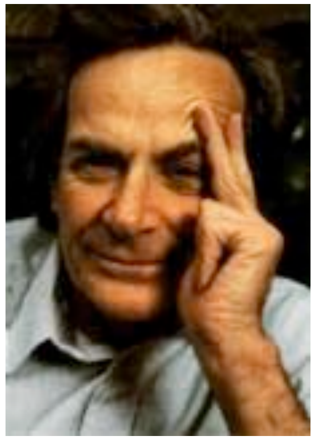
Noise regulates control authority:

$$\lambda \mathbf{G}(\mathbf{x})\mathbf{R}^{-1}\mathbf{G}(\mathbf{x})^T = \mathbf{B}(\mathbf{x})\mathbf{B}(\mathbf{x})^T$$

Statistical Physics

Chapman Kolmogorov:

$$-\partial_t \Psi = -\frac{1}{\lambda}\tilde{q}\Psi + \tilde{\mathbf{f}}^T(\nabla_{\mathbf{x}}\Psi) + \frac{1}{2} tr\left((\nabla_{\mathbf{xx}}\Psi)\mathbf{B}\mathbf{B}^T\right)$$



Richard Feynman

Feynman - Kac

Sample from the uncontrolled dynamics: $d\mathbf{x} = \tilde{\mathbf{f}}(\mathbf{x})dt + \mathbf{B}(\mathbf{x})d\omega$

Evaluate the expectation: $\Psi(\mathbf{x}, t_i) = \mathbb{E}_{\mathbb{P}}\left[\exp\left(-\frac{1}{\lambda}\int_{t_i}^{t_N}\tilde{q}(\mathbf{x})dt\right)\Psi(\mathbf{x}, t_N)\right]$

Find optimal Controls: $\mathbf{u}_{PI}(\mathbf{x}) = -\mathbf{R}^{-1}\left(q_1(\mathbf{x}, t) - \lambda\mathbf{G}(\mathbf{x})^T\frac{\nabla_{\mathbf{x}}\Psi(\mathbf{x}, t)}{\Psi(\mathbf{x}, t)}\right)$

Push the dynamics to states with high desirability.

Mark Kac

Making Control Linear

Hamilton Jacobi Bellman:

$$-\partial_t V = \tilde{q} + (\nabla_{\mathbf{x}} V)^T \tilde{\mathbf{f}} - \frac{1}{2}(\nabla_{\mathbf{x}} V)^T \mathbf{G}\mathbf{R}^{-1}\mathbf{G}^T(\nabla_{\mathbf{x}} V) + \frac{1}{2} tr\left((\nabla_{\mathbf{xx}} V)\mathbf{B}\mathbf{B}^T\right)$$

Desirability:

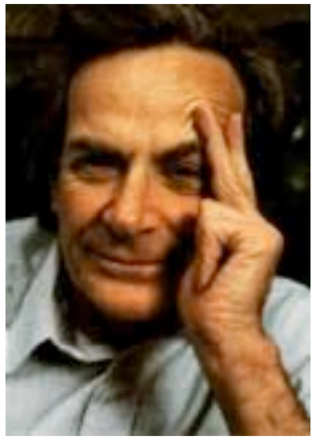
$$\Psi(\mathbf{x}, t) = \exp\left(-\frac{1}{\lambda} V(\mathbf{x}, t)\right)$$

Noise regulates control authority:

$$\lambda \mathbf{G}(\mathbf{x})\mathbf{R}^{-1}\mathbf{G}(\mathbf{x})^T = \mathbf{B}(\mathbf{x})\mathbf{B}(\mathbf{x})^T$$

Statistical Physics

Chapman Kolmogorov: $$-\partial_t \Psi = -\frac{1}{\lambda}\tilde{q}\Psi + \tilde{\mathbf{f}}^T(\nabla_{\mathbf{x}}\Psi) + \frac{1}{2} tr\left((\nabla_{\mathbf{xx}}\Psi)\mathbf{B}\mathbf{B}^T\right)$$



Richard Feynman

Feynman - Kac

Sample from the uncontrolled dynamics: $\mathrm{d}\mathbf{x} = \tilde{\mathbf{f}}(\mathbf{x})\mathrm{d}t + \mathbf{B}(\mathbf{x})\mathrm{d}\omega$

Evaluate the expectation: $$\Psi(\mathbf{x}, t_i) = \mathbb{E}_{\mathbb{P}}\left[\exp\left(-\frac{1}{\lambda}\int_{t_i}^{t_N} \tilde{q}(\mathbf{x})\mathrm{d}t\right)\Psi(\mathbf{x}, t_N)\right]$$

Mark Kac

Making Control Linear

Hamilton Jacobi Bellman:

$$-\partial_t V = \tilde{q} + (\nabla_{\mathbf{x}} V)^T \tilde{\mathbf{f}} - \frac{1}{2} (\nabla_{\mathbf{x}} V)^T \mathbf{G} \mathbf{R}^{-1} \mathbf{G}^T (\nabla_{\mathbf{x}} V) + \frac{1}{2} tr\left((\nabla_{\mathbf{xx}} V) \mathbf{B} \mathbf{B}^T \right)$$

Desirability:

$$\Psi(\mathbf{x}, t) = \exp\left(-\frac{1}{\lambda} V(\mathbf{x}, t) \right)$$

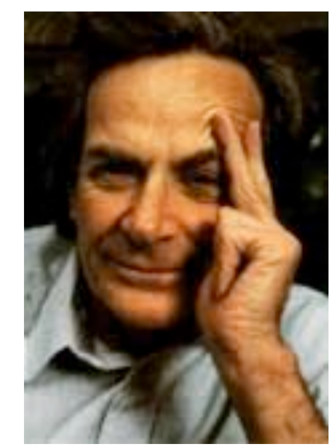
Noise regulates control authority:

$$\lambda \mathbf{G}(\mathbf{x}) \mathbf{R}^{-1} \mathbf{G}(\mathbf{x})^T = \mathbf{B}(\mathbf{x}) \mathbf{B}(\mathbf{x})^T$$

Statistical Physics

Chapman Kolmogorov:

$$-\partial_t \Psi = -\frac{1}{\lambda} \tilde{q} \Psi + \tilde{\mathbf{f}}^T (\nabla_{\mathbf{x}} \Psi) + \frac{1}{2} tr\left((\nabla_{\mathbf{xx}} \Psi) \mathbf{B} \mathbf{B}^T \right)$$



Richard Feynman

Feynman - Kac

Sample from the uncontrolled dynamics: $\mathrm{d}\mathbf{x} = \tilde{\mathbf{f}}(\mathbf{x}) \mathrm{d}t + \mathbf{B}(\mathbf{x}) \mathrm{d}\omega$

Evaluate the expectation: $\Psi(\mathbf{x}, t_i) = \mathbb{E}_{\mathbb{P}} \left[\exp\left(-\frac{1}{\lambda} \int_{t_i}^{t_N} \tilde{q}(\mathbf{x}) \mathrm{d}t \right) \Psi(\mathbf{x}, t_N) \right]$

Lower bound: $V(\mathbf{x}, t_i) = -\lambda \log \mathbb{E}_{\mathbb{P}} \left[\exp\left(-\frac{1}{\lambda} \int_{t_i}^{t_N} \tilde{q}(\mathbf{x}, t) \mathrm{d}t \right) \Psi(\mathbf{x}, t_N) \right]$

$$\leq \mathbb{E}_{\mathbb{Q}} \left[\phi(\mathbf{x}, t_N) + \int_{t_i}^{t_N} \mathcal{L}(\mathbf{x}, \mathbf{u}, t) \mathrm{d}t \right]$$

Mark Kac

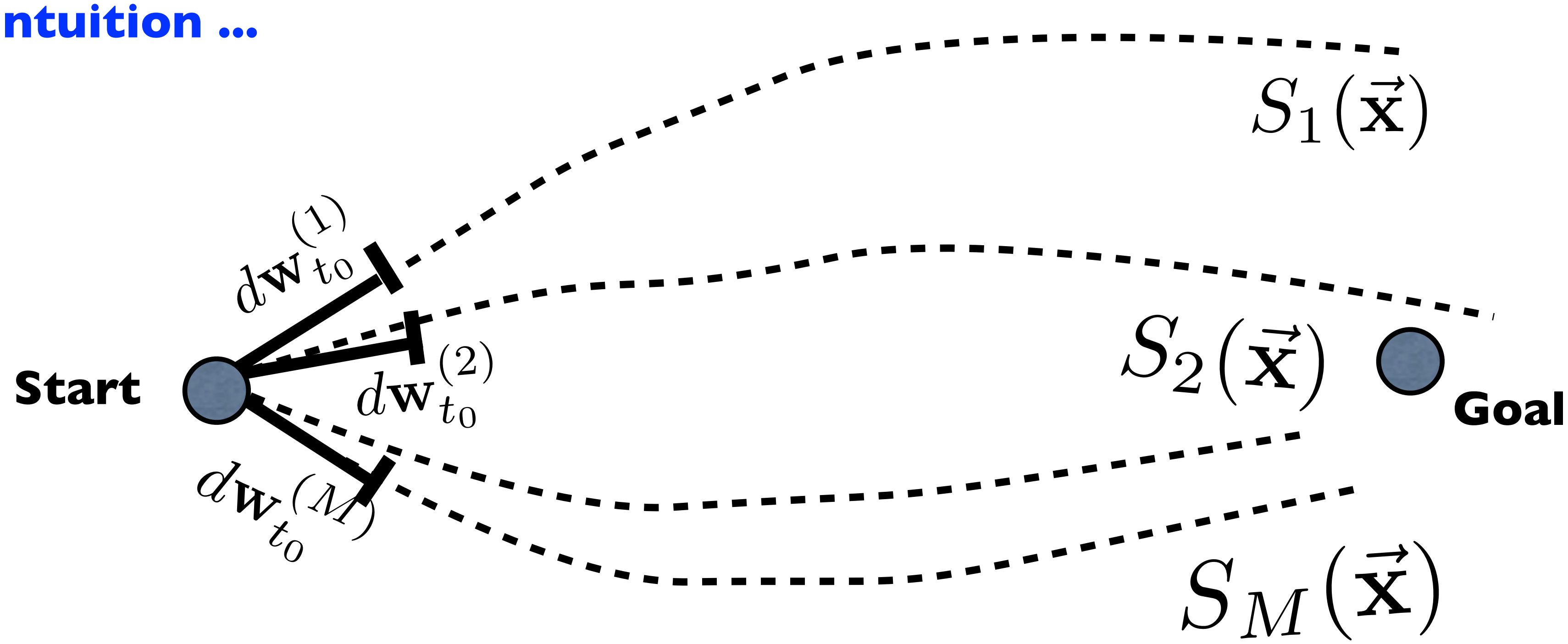
Generalized Path Integral Control

Path Integral Control: $\mathbf{u}(\mathbf{x}_{t_i}, t_i)dt = -\mathbf{R}^{-1} q_1(\mathbf{x}_{t_i}, t_i)dt + \int \mathcal{P}(\vec{\mathbf{x}})\mathbf{u}_L(\mathbf{x}_{t_i}, t_i)d\vec{\mathbf{x}}$

Local Controls: $\mathbf{u}_L(\mathbf{x}_{t_i}, t_i) = \underbrace{\mathbf{R}^{-1}\mathbf{G}(\mathbf{x}_{t_i})^T \left(\mathbf{G}(\mathbf{x}_{t_i})\mathbf{R}^{-1}\mathbf{G}(\mathbf{x}_{t_i})^T\right)^{-1}}_{\mathbf{Y}} \mathbf{G}(\mathbf{x}_{t_i})d\mathbf{w}(t_i)$

Trajectory Probability: $\mathcal{P}(\vec{\mathbf{x}}) = \frac{\exp\left(-\frac{1}{\lambda}S(\vec{\mathbf{x}})\right)}{\int \exp\left(-\frac{1}{\lambda}S(\vec{\mathbf{x}})\right)d\vec{\mathbf{x}}}$

Intuition ...



$$\mathbf{u} = \sum_i P_i \; d\mathbf{w}_{t_o}^{(i)}$$

$$P_i = \frac{\exp(-S_i(\vec{\mathbf{x}}))}{\sum_i \exp(-S_i(\vec{\mathbf{x}}))}$$

Sample from the uncontrolled dynamics:

E.Theodorou, J. Buchli ,S. Schaal. AISTATS2009, JMLR2010

Iterative Path Integral Control

Sampling with uncontrolled dynamics:

$$\Psi(\mathbf{x}, t_i) = \int \exp\left(-\frac{1}{\lambda}\tilde{q}(\mathbf{x}_t, t)dt \right) \Psi(\mathbf{x}, t_N) d\mathbb{Q}_{\textbf{Uncontrolled}}$$

Sampling with controlled dynamics:

$$\Psi(\mathbf{x}, t_i) = \int \exp\left(-\frac{1}{\lambda}\tilde{q}(\mathbf{x}_t, t)dt \right) \Psi(\mathbf{x}, t_N) \boxed{\frac{d\mathbb{Q}_{\textbf{Uncontrolled}}}{d\mathbb{P}_{\textbf{Controlled}}}} d\mathbb{P}_{\textbf{Controlled}}$$

Radon Nikodym derivative

Find optimal Controls: $\mathbf{u}_{PI}(\mathbf{x}) = -\mathbf{R}^{-1}\left(q_1(\mathbf{x}, t) - \lambda \mathbf{G}(\mathbf{x})^T \frac{\nabla_{\mathbf{x}} \Psi(\mathbf{x}, t)}{\Psi(\mathbf{x}, t)} \right)$

Iterative Path Integral Control: $\mathbf{u}_{new}(t_k) = \mathbf{u}_{old}(t_k) + \sum_i P_i \; d\mathbf{w}^{(i)}(t_k)$

Path Probability: $P_i = \frac{\exp\left(- S(\vec{\mathbf{x}}_i) + \text{correction} \right)}{\sum_i \exp\left(- S(\vec{\mathbf{x}}_i) + \text{correction} \right)}$

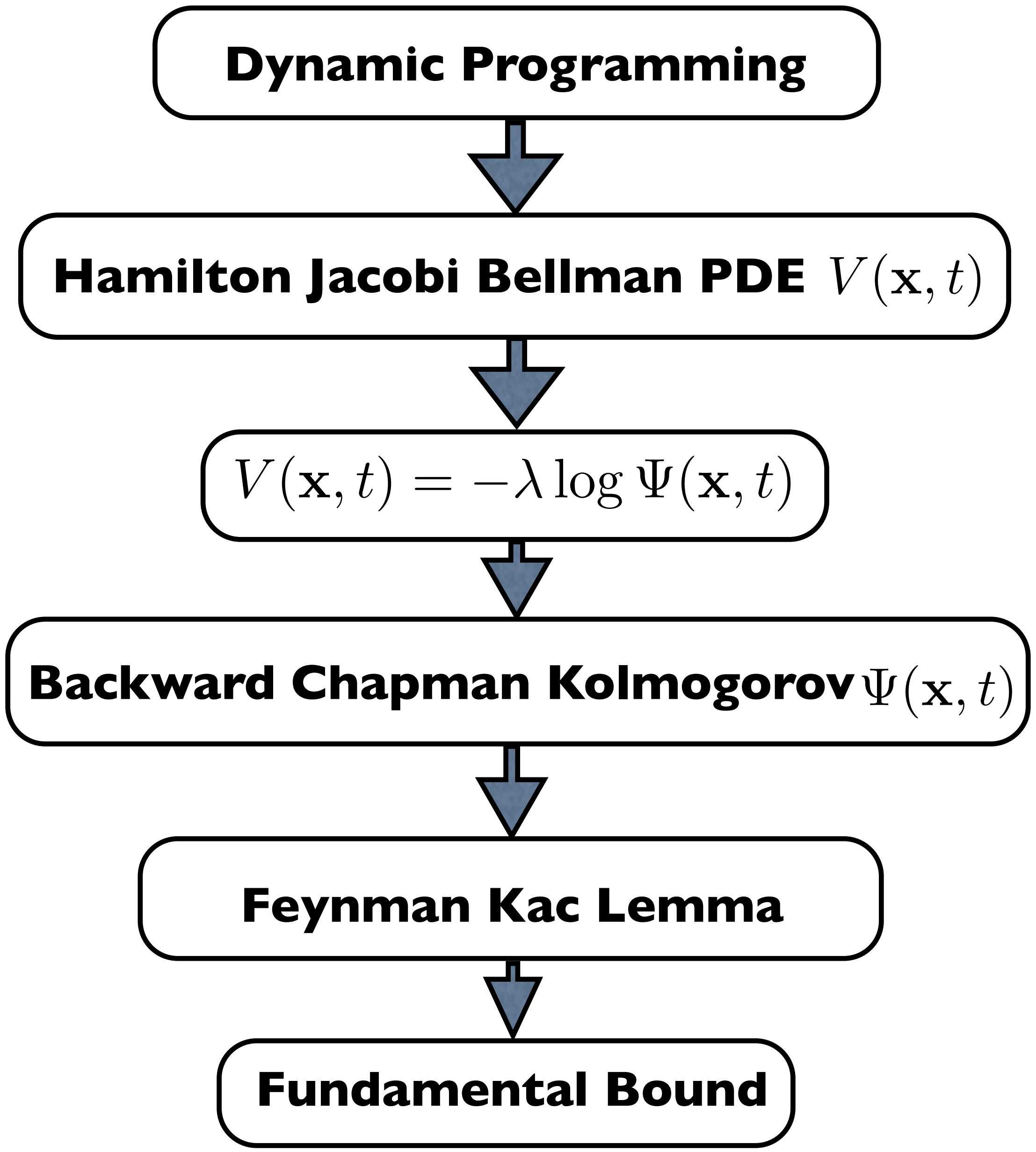
E.Theodorou, J. Buchli ,S. Schaal JMLR2010

E.Theodorou, E. Todorov. CDC2012

Overview

Optimal Control Theory

Constrained Optimization



E. Theodorou, E Todorov. CDC 2012

E. Theodorou, Entropy 2015

Information Theoretic View

Free Energy: $\mathcal{F} = -\frac{1}{|\rho|} \log \int_{\Omega} \exp\Big(-|\rho| J(\mathbf{x}) \Big) \mathrm{d}\mathbb{P}(\omega)$

Generalized Entropy: $S(\mathbb{Q}||\mathbb{P}) = -\int_{\Omega} \frac{\mathrm{d}\mathbb{Q}(\omega)}{\mathrm{d}\mathbb{P}(\omega)} \log \frac{\mathrm{d}\mathbb{Q}(\omega)}{\mathrm{d}\mathbb{P}(\omega)} \mathrm{d}\mathbb{P}(\omega)$

$$-\frac{1}{|\rho|} \log \int \exp\Big(-|\rho| J(\mathbf{x}) \Big) \mathrm{d}\mathbb{P}(\omega) \leq \mathbb{E}_{\mathbb{Q}}\Big[J(\mathbf{x}) \Big] + \frac{1}{|\rho|} \mathbb{KL}\Big(\mathbb{Q}||\mathbb{P} \Big)$$

$$\textbf{Free Energy} \leq \textbf{Work} - \textbf{Temperature} \cdot \textbf{Generalized Entropy}$$

Optimal Measure: $\mathrm{d}\mathbb{Q}^* = \frac{\exp\Big(-|\rho| J(\mathbf{x}) \Big) \mathrm{d}\mathbb{P}}{\int \exp\Big(-|\rho| J(\mathbf{x}) \Big) \mathrm{d}\mathbb{P}}$

Non-Classical View

Uncontrolled dynamics: $\mathbb{P} : \mathrm{d}\mathbf{x} = f(\mathbf{x})\mathrm{d}t + \frac{1}{\sqrt{|\rho|}}\boldsymbol{\mathcal{B}}(\mathbf{x})\mathrm{d}\omega^{(0)}$

Controlled dynamics: $\mathbb{Q} : \mathrm{d}\mathbf{x} = f(\mathbf{x})\mathrm{d}t + \boldsymbol{\mathcal{B}}(\mathbf{x})\left(\mathbf{u}\mathrm{d}t + \frac{1}{\sqrt{|\rho|}}\mathrm{d}\omega^{(1)}\right)$

Fundamental Relationship: $\xi = -\frac{1}{|\rho|}\log \mathbb{E}_{\mathbb{P}}\left[\exp\left(-|\rho|\mathcal{J}(\mathbf{x})\right)\right] \leq \mathbb{E}_{\mathbb{Q}}\left[\mathcal{J}(\mathbf{x})\right] + \frac{1}{|\rho|}\mathbb{KL}\left(\mathbb{Q}\middle|\middle|\mathbb{P}\right)$

Radon Nikodym: $\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\frac{1}{2}|\rho|\int_{t_i}^{t_N}\mathbf{u}^T\mathbf{u}dt + \sqrt{|\rho|}\int_{t_i}^{t_N}\mathbf{u}^T d\mathbf{w}^{(1)}(t)\right)$

Fundamental Relationship: $\xi(\mathbf{x}, t_i) = -\frac{1}{|\rho|}\log \mathbb{E}_{\mathbb{P}}\left[\exp\left(-|\rho|\mathcal{J}(\mathbf{x})\right)\right] \leq \mathbb{E}_{\mathbb{Q}}\left[\mathcal{J}(\mathbf{x}) + \int_{t_i}^{t_N}\frac{1}{2}\mathbf{u}^{\mathrm{T}}\mathbf{u}\mathrm{d}t\right]$

Free Energy **Cost Function**

- ✦**There is a lower bound in the cost function**
- ✦**How to find this lower bound**
- ✦**Permits generalizations to other classes of stochastic dynamics.**

Non-Classical View

Basic Relationship: $\xi(\mathbf{x}, t_i) = -\frac{1}{|\rho|} \log \underbrace{\int \exp\Big(-|\rho| J(\mathbf{x})\Big) \mathrm{d}\mathbb{P}(\omega)}_{\Phi(\mathbf{x},t)} \leq \mathbb{E}_{\mathbb{Q}}\Big[J(\mathbf{x}) + \int_{t_i}^{t_N} \frac{1}{2}\mathbf{u}^{\mathrm{T}}\mathbf{R}\mathbf{u}\mathrm{d}t\Big]$

Backward Chapman Kolmogorov:

$$-\partial_t \Phi = -|\rho| q_0 \Phi + \mathbf{f}^T(\nabla_{\mathbf{x}}\Phi) + \frac{1}{2|\rho|} tr\left((\nabla_{\mathbf{xx}}\Phi)\mathbf{B}\mathbf{B}^T\right)$$

Exponential Transformation: $\Phi(\mathbf{x}, t) = \exp\left(-|\rho|\xi(\mathbf{x}, t)\right)$

Hamilton Jacobi Bellman-PDE:

$$\boxed{-\partial_t \xi = q_0 + (\nabla_{\mathbf{x}}\xi)^T \mathbf{f} - \underline{\frac{1}{2}(\nabla_{\mathbf{x}}\xi)^T \mathbf{B}\mathbf{B}^T (\nabla_{\mathbf{x}}\xi)} + \frac{1}{2|\rho|} tr\left((\nabla_{\mathbf{xx}}\xi)\mathbf{B}\mathbf{B}^T\right)}$$

$\xi(\mathbf{x}, t)$ **: is a Value function**

$\Phi(\mathbf{x}, t)$ **: is a desirability function**

Overview

Optimal Control Theory

Constrained Optimization

- Dynamic Programming
- ↓
- Hamilton Jacobi Bellman PDE $V(\mathbf{x}, t)$
- ↓
- $V(\mathbf{x}, t) = -\lambda \log \Psi(\mathbf{x}, t)$
- ↓
- Backward Chapman Kolmogorov $\Psi(\mathbf{x}, t)$
- ↓
- Feynman Kac Lemma
- ↓
- Fundamental Bound

$$\Psi(\mathbf{x}, t) = \Phi(\mathbf{x}, t)$$

$$V(\mathbf{x}, t) = \xi(\mathbf{x}, t)$$

Information Theory

Free Energy & Relative Entropy

- Jensens Inequality & Girsanov Theorem
- ↓
- Fundamental Bound
- ↓
- Feynman Kac Lemma
- ↓
- Backward Chapman Kolmogorov $\Phi(\mathbf{x}, t)$
- ↓
- $\xi(\mathbf{x}, t) = -\frac{1}{\rho} \log \Phi(\mathbf{x}, t)$
- ↓
- Hamilton Jacobi Bellman PDE

E. Theodorou, E Todorov, CDC 2012.

E. Theodorou, Entropy 2015.

Non-Classical View

Fundamental Relationship: $$-\frac{1}{|\rho|}\log\int \exp\Big(-|\rho| J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}(\omega) \leq \mathbb{E}_{\mathbb{Q}}\Big[J(\mathbf{x})\Big] + \frac{1}{|\rho|}\mathbb{KL}\Big(\mathbb{Q}||\mathbb{P}\Big)$$

Uncontrolled dynamics: $$\mathbb{P} : \mathrm{d}\mathbf{x} = \mathbf{F}(\mathbf{x}, 0)\mathrm{d}t + \mathbf{C}(\mathbf{x})\mathrm{d}\mathbf{w}^{(0)}$$

Controlled dynamics: $$\mathbb{Q} : \mathrm{d}\mathbf{x} = \mathbf{F}(\mathbf{x}, \mathbf{u})\mathrm{d}t + \mathbf{C}(\mathbf{x})\mathrm{d}\mathbf{w}^{(1)}$$

Grady et all, ICRA 2016.

Non-Classical View

Fundamental Relationship: $$-\frac{1}{|\rho|}\log\int\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}(\omega) \leq \mathbb{E}_{\mathbb{Q}}\Big[J(\mathbf{x})\Big] + \frac{1}{|\rho|}\mathbb{KL}\Big(\mathbb{Q}||\mathbb{P}\Big)$$

Uncontrolled dynamics: $$\mathbb{P} : \mathrm{d}\mathbf{x} = \mathbf{F}(\mathbf{x}, 0)\mathrm{d}t + \mathbf{C}(\mathbf{x})\mathrm{d}\mathbf{w}^{(0)}$$

Controlled dynamics: $$\mathbb{Q} : \mathrm{d}\mathbf{x} = \mathbf{F}(\mathbf{x}, \mathbf{u})\mathrm{d}t + \mathbf{C}(\mathbf{x})\mathrm{d}\mathbf{w}^{(1)}$$

Optimal Measure:

$$\mathrm{d}\mathbb{Q}^* = \frac{\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}}{\int\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}}$$

New Optimal Control Formulation:

$$\mathbf{u} = \mathrm{argmin}\ \mathbb{KL}\Big(\mathbb{Q}^*||\mathbb{Q}(\mathbf{u})\Big)$$

Grady et all, ICRA 2016.

Non-Classical View

Fundamental Relationship: $$-\frac{1}{|\rho|}\log\int\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}(\omega) \leq \mathbb{E}_{\mathbb{Q}}\Big[J(\mathbf{x})\Big] + \frac{1}{|\rho|}\mathbb{KL}\Big(\mathbb{Q}||\mathbb{P}\Big)$$

Uncontrolled dynamics: $$\mathbb{P}: \mathrm{d}\mathbf{x} = \mathbf{F}(\mathbf{x},0)\mathrm{d}t + \mathbf{C}(\mathbf{x})\mathrm{d}\mathbf{w}^{(0)}$$

Controlled dynamics: $$\mathbb{Q}: \mathrm{d}\mathbf{x} = \mathbf{F}(\mathbf{x},\mathbf{u})\mathrm{d}t + \mathbf{C}(\mathbf{x})\mathrm{d}\mathbf{w}^{(1)}$$

Optimal Measure:

$$\mathrm{d}\mathbb{Q}^* = \frac{\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}}{\int\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}}$$

New Optimal Control Formulation:

$$\mathbf{u} = \operatorname{argmin}\ \mathbb{KL}\Big(\mathbb{Q}^*||\mathbb{Q}(\mathbf{u})\Big)$$

Optimal Measure: $$\frac{\mathrm{d}\mathbb{Q}^*}{\mathrm{d}\mathbb{P}}$$

Radon Nikodym: $$\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}(\mathbf{u})}$$

Grady et all, ICRA 2016.

Non-Classical View

Fundamental Relationship: $$-\frac{1}{|\rho|}\log\int\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}(\omega) \le \mathbb{E}_{\mathbb{Q}}\Big[J(\mathbf{x})\Big] + \frac{1}{|\rho|}\mathbb{KL}\Big(\mathbb{Q}||\mathbb{P}\Big)$$

Uncontrolled dynamics: $$\mathbb{P} : \mathrm{d}\mathbf{x} = \mathbf{F}(\mathbf{x}, 0)\mathrm{d}t + \mathbf{C}(\mathbf{x})\mathrm{d}\mathbf{w}^{(0)}$$

Controlled dynamics: $$\mathbb{Q} : \mathrm{d}\mathbf{x} = \mathbf{F}(\mathbf{x}, \mathbf{u})\mathrm{d}t + \mathbf{C}(\mathbf{x})\mathrm{d}\mathbf{w}^{(1)}$$

Optimal Measure:

$$\mathrm{d}\mathbb{Q}^* = \frac{\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}}{\int\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}}$$

New Optimal Control Formulation:

$$\mathbf{u} = \mathrm{argmin}\ \mathbb{KL}\Big(\mathbb{Q}^*||\mathbb{Q}(\mathbf{u})\Big)$$

Grady et all, ICRA 2016.

Non-Classical View

Fundamental Relationship: $$-\frac{1}{|\rho|}\log\int\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}(\omega) \leq \mathbb{E}_{\mathbb{Q}}\Big[J(\mathbf{x})\Big] + \frac{1}{|\rho|}\mathbb{KL}\Big(\mathbb{Q}||\mathbb{P}\Big)$$

Uncontrolled dynamics: $$\mathbb{P} : \mathrm{d}\mathbf{x} = \mathbf{F}(\mathbf{x}, 0)\mathrm{d}t + \mathbf{C}(\mathbf{x})\mathrm{d}\mathbf{w}^{(0)}$$

Controlled dynamics: $$\mathbb{Q} : \mathrm{d}\mathbf{x} = \mathbf{F}(\mathbf{x}, \mathbf{u})\mathrm{d}t + \mathbf{C}(\mathbf{x})\mathrm{d}\mathbf{w}^{(1)}$$

Optimal Measure:

$$\mathrm{d}\mathbb{Q}^* = \frac{\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}}{\int\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}}$$

New Optimal Control Formulation:

$$\mathbf{u} = \operatorname{argmin}\ \mathbb{KL}\Big(\mathbb{Q}^*||\mathbb{Q}(\mathbf{u})\Big)$$

Control parameterization:

$$\mathbf{u}_t = \begin{cases} \mathbf{u}_0 & \text{if} \quad 0 \leq t < \Delta t \\ \mathbf{u}_1 & \text{if} \quad \Delta t \leq t < 2\Delta t \\ \vdots & \\ \mathbf{u}_j & \text{if } j\Delta t \leq t < (j+1)\Delta t \\ \vdots & \end{cases}$$

Grady et all, ICRA 2016.

Non-Classical View

Fundamental Relationship: $$-\frac{1}{|\rho|}\log\int\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}(\omega) \leq \mathbb{E}_{\mathbb{Q}}\Big[J(\mathbf{x})\Big] + \frac{1}{|\rho|}\mathbb{KL}\Big(\mathbb{Q}||\mathbb{P}\Big)$$

Uncontrolled dynamics: $$\mathbb{P} : \mathrm{d}\mathbf{x} = \mathbf{F}(\mathbf{x}, 0)\mathrm{d}t + \mathbf{C}(\mathbf{x})\mathrm{d}\mathbf{w}^{(0)}$$

Controlled dynamics: $$\mathbb{Q} : \mathrm{d}\mathbf{x} = \mathbf{F}(\mathbf{x}, \mathbf{u})\mathrm{d}t + \mathbf{C}(\mathbf{x})\mathrm{d}\mathbf{w}^{(1)}$$

Optimal Measure:

$$\mathrm{d}\mathbb{Q}^* = \frac{\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}}{\int\exp\Big(-|\rho|J(\mathbf{x})\Big)\mathrm{d}\mathbb{P}}$$

New Optimal Control Formulation:

$$\mathbf{u} = \mathrm{argmin}\ \mathbb{KL}\Big(\mathbb{Q}^*||\mathbb{Q}(\mathbf{u})\Big)$$

Control parameterization:

$$\mathbf{u}_t = \begin{cases} \mathbf{u}_0 & \text{if} \quad 0 \leq t < \Delta t \\ \mathbf{u}_1 & \text{if} \quad \Delta t \leq t < 2\Delta t \\ \vdots & \\ \mathbf{u}_j & \text{if } j\Delta t \leq t < (j+1)\Delta t \\ \vdots & \end{cases}$$

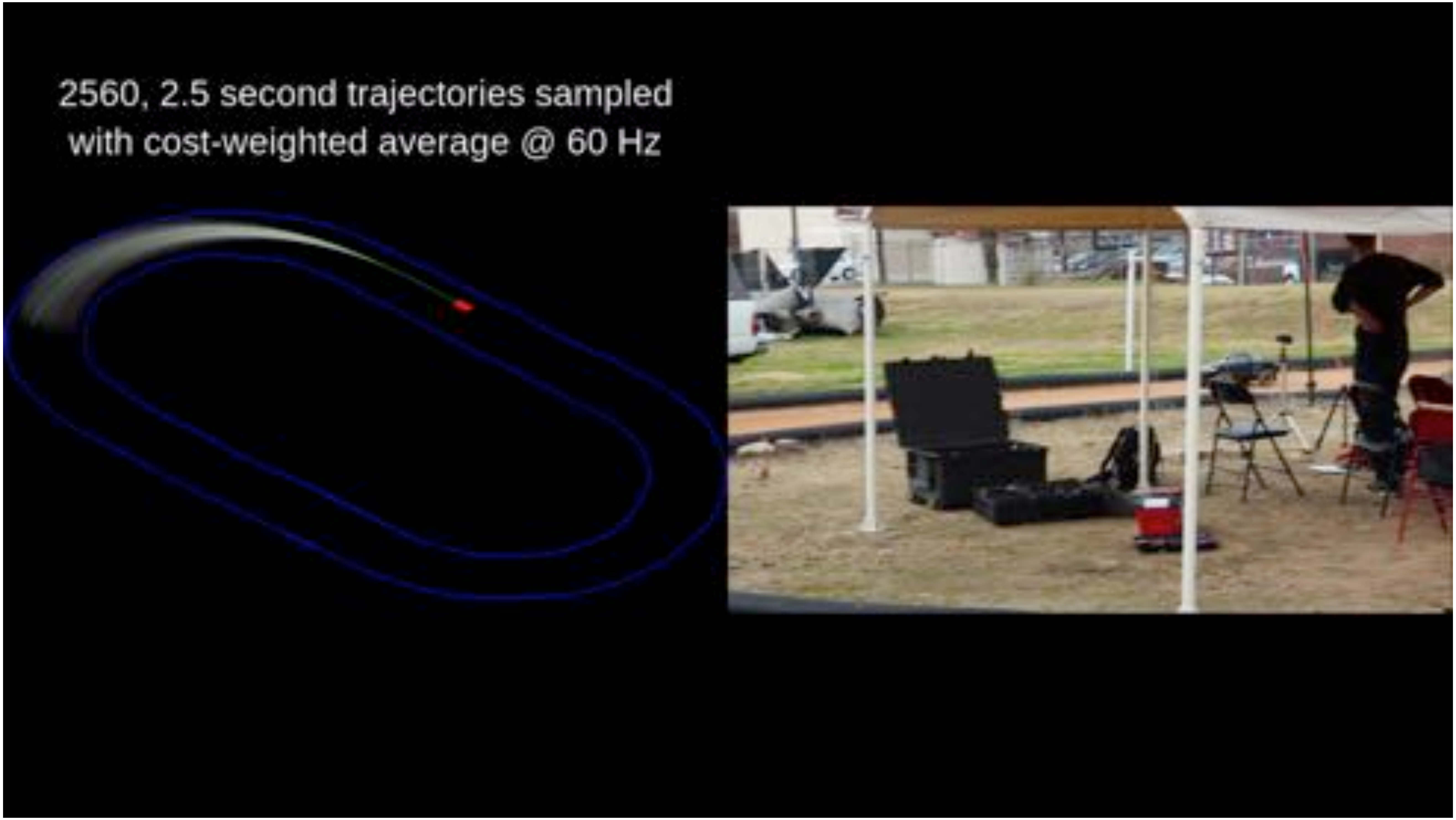
Generalized Importance Sampling:

$$\mathbb{E}_{\mathbb{P}}\Big[f(\mathbf{x})\Big] = \mathbb{E}_{\mathbb{Q}_{m,\Sigma}}\Big[f(\mathbf{x})\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}_{m,\Sigma}}\Big]$$

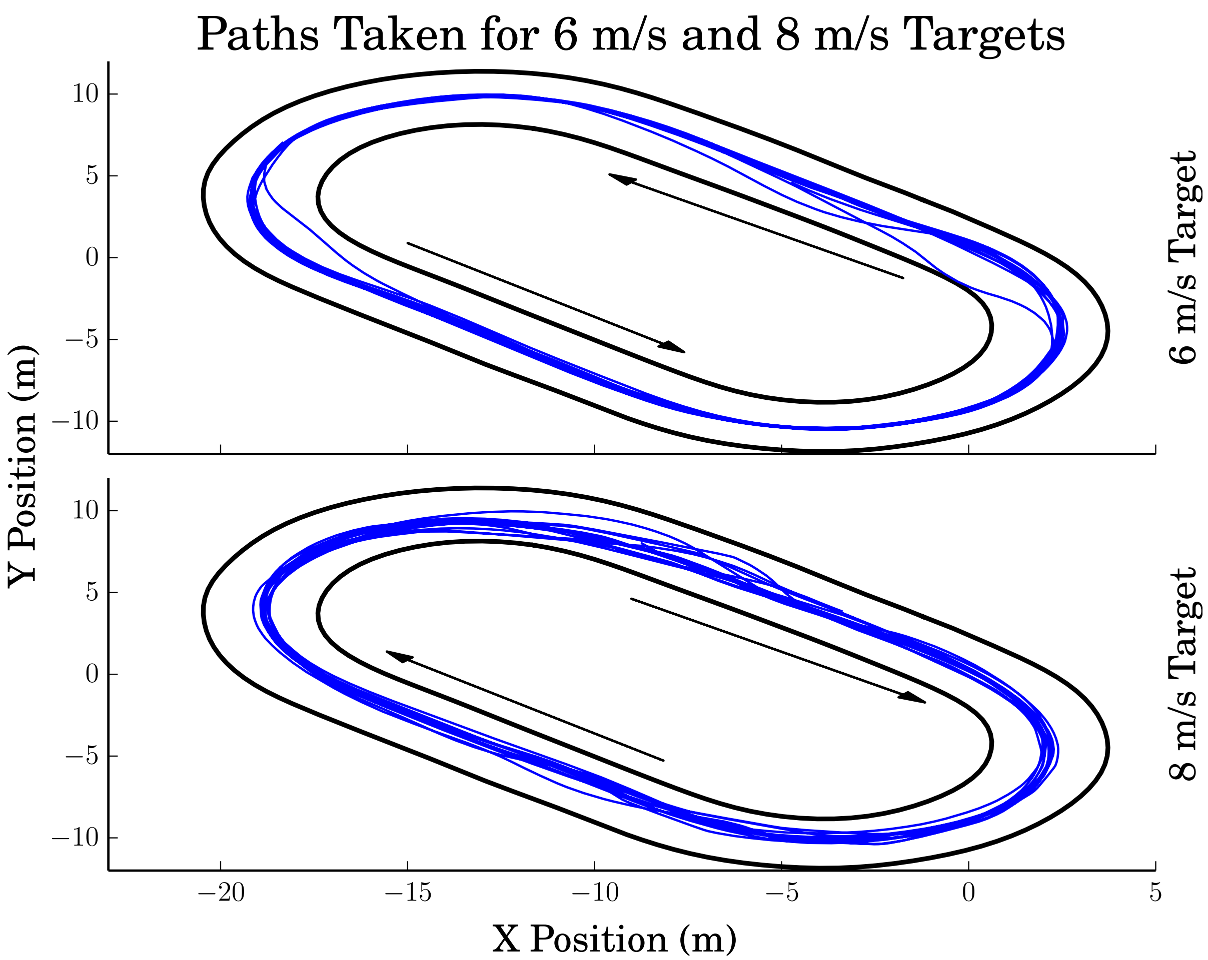
Grady et all, ICRA 2016.

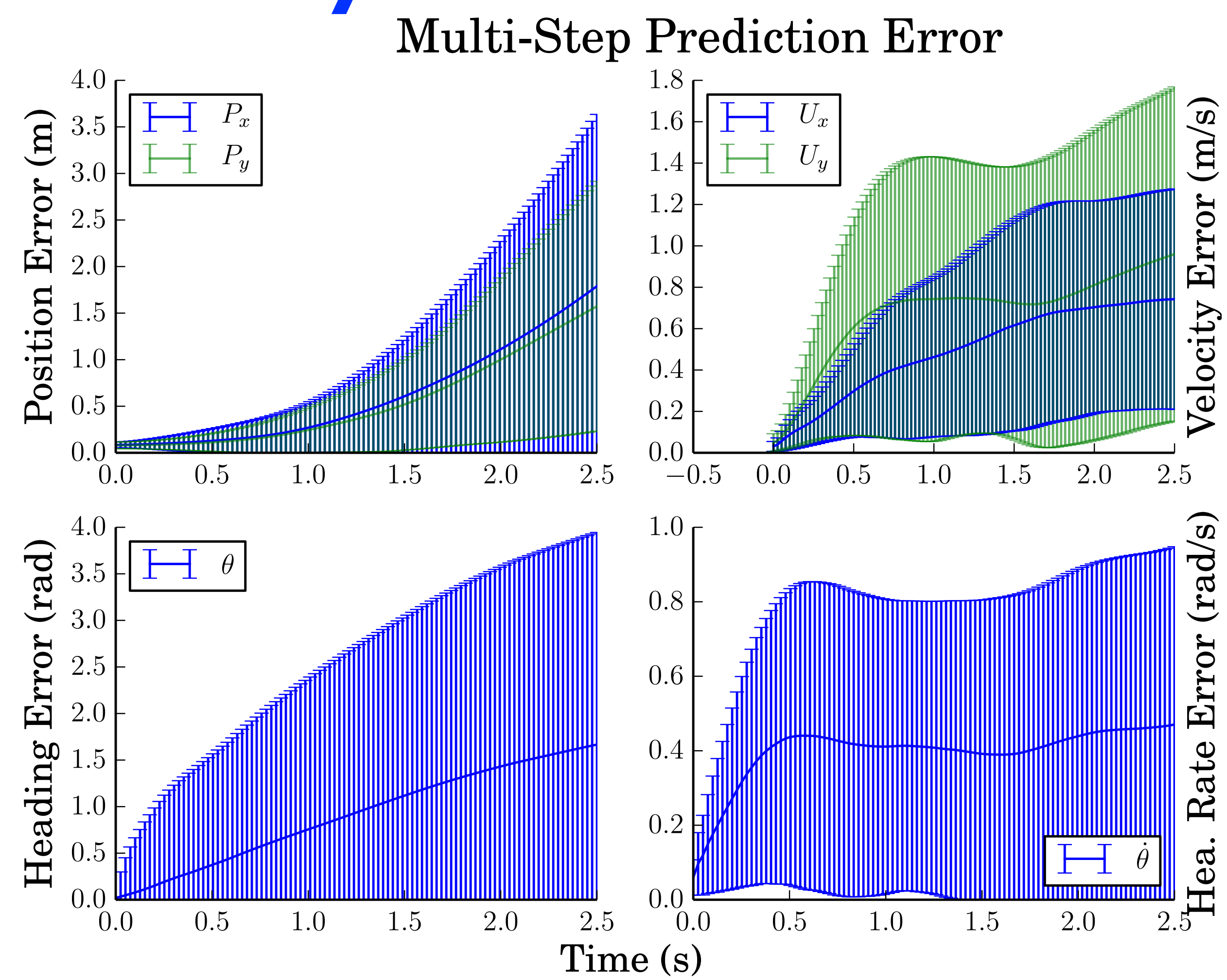
MPPI with offline model learning

7~8 m/sec



Applications in Robotics-High Speed MPPI with offline learned dynamics





Controlled dynamics: $f_i(\mathbf{x}) \sim \mathcal{N}(\mu_i(\mathbf{x}), \sigma_i^2(\mathbf{x}))$

Mean: $\mu_i(\mathbf{x}) = \frac{1}{\sigma_n^2}\phi(\tilde{\mathbf{x}})^{\mathrm{T}} A_i^{-1} \Phi(\tilde{\mathbf{X}}) Y_i$

Variance: $\sigma_i^2(\tilde{\mathbf{x}}) = \phi(\tilde{\mathbf{x}})^{\mathrm{T}} A_i^{-1} \phi(\tilde{\mathbf{x}})$

$$A_i = \frac{1}{\sigma_n^2}\Phi(\tilde{\mathbf{x}})\Phi(\tilde{\mathbf{x}})^{\mathrm{T}} + \Sigma_p^{-1}$$

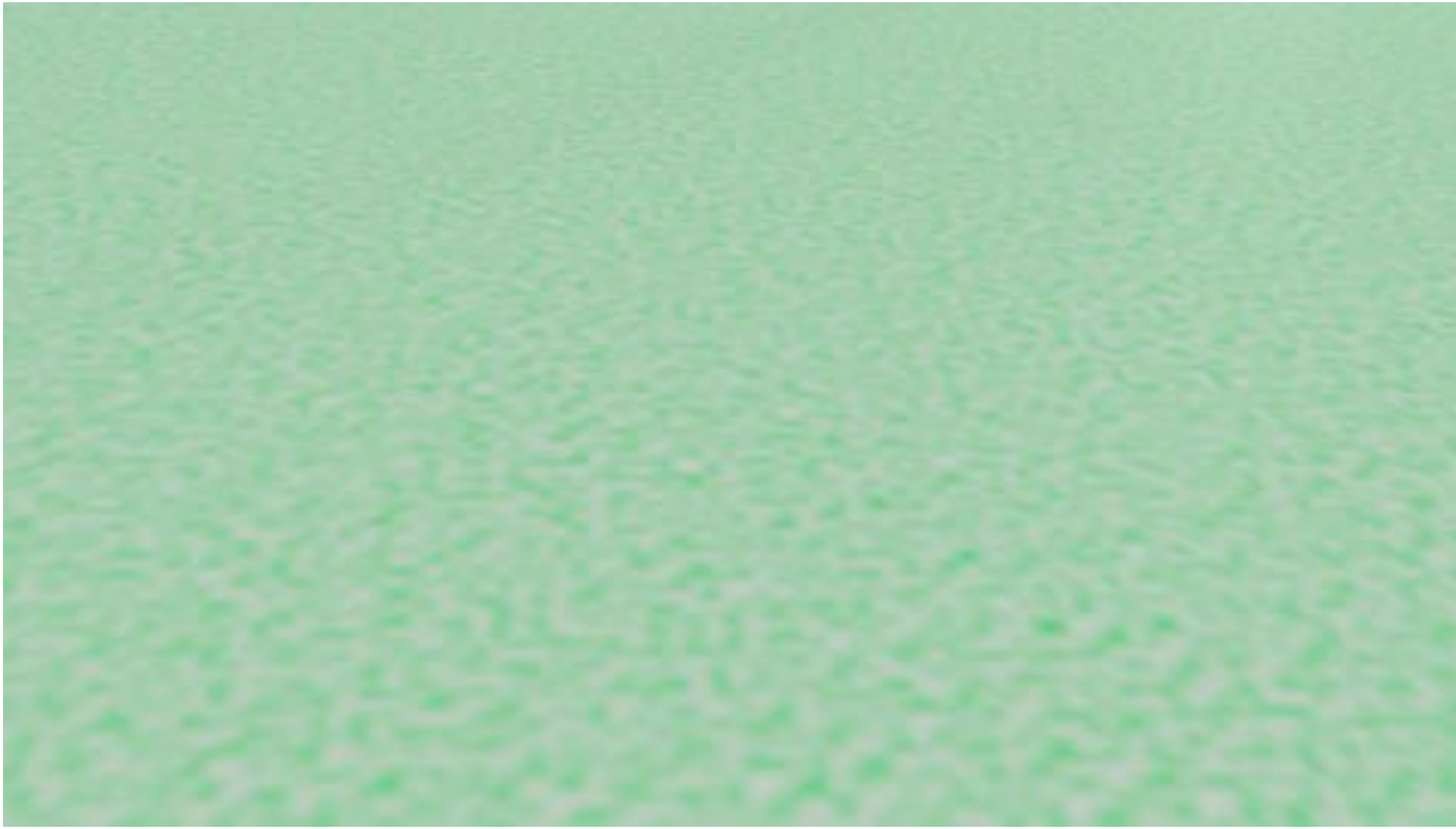
Applications in Robotics-MPPI with online model learning

8~9 m/sec

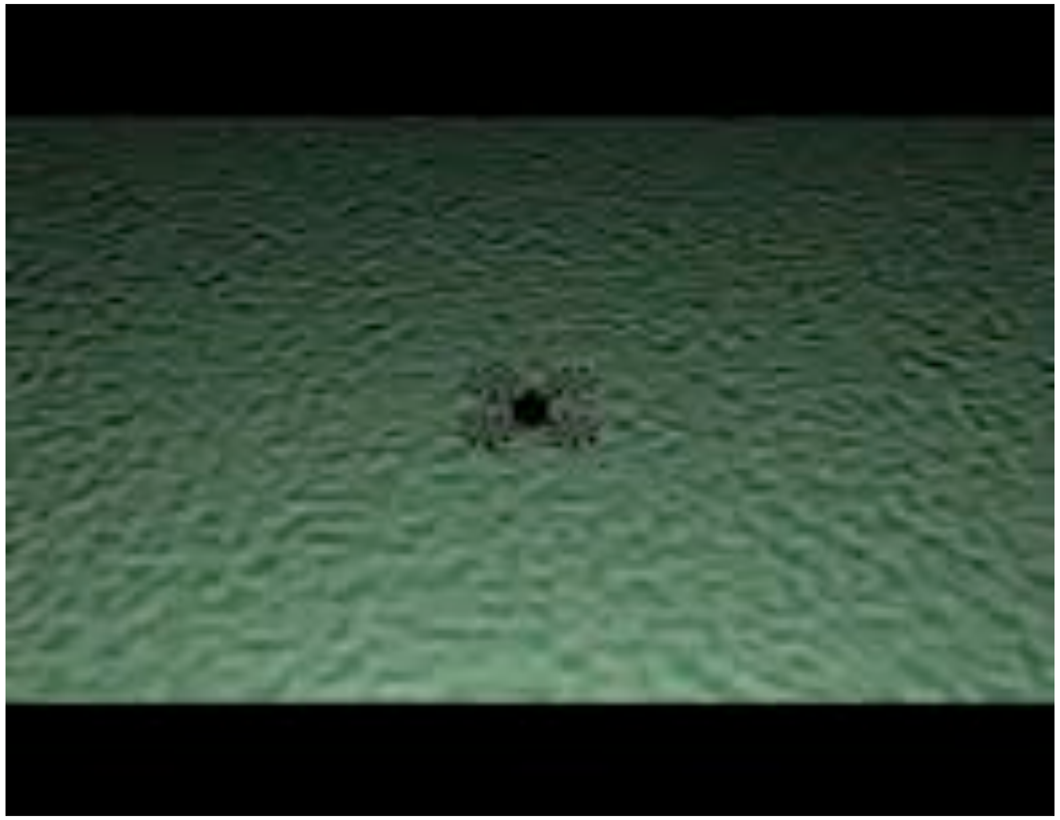
Model Predictive Path Integral Control using Artificial Neural Networks

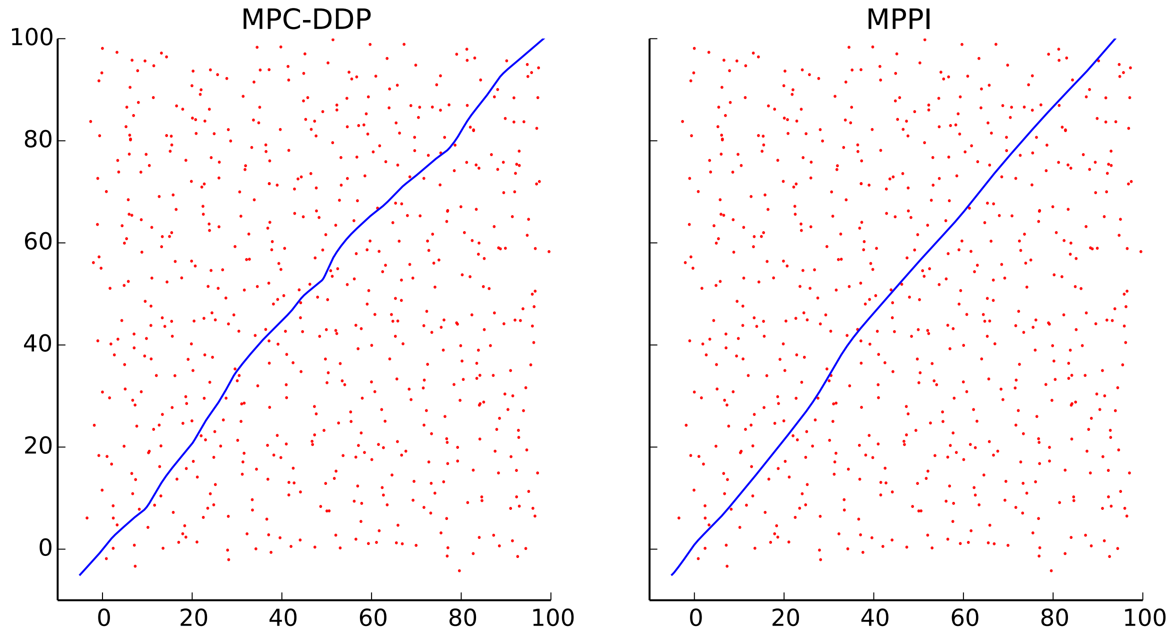
~11 m/sec

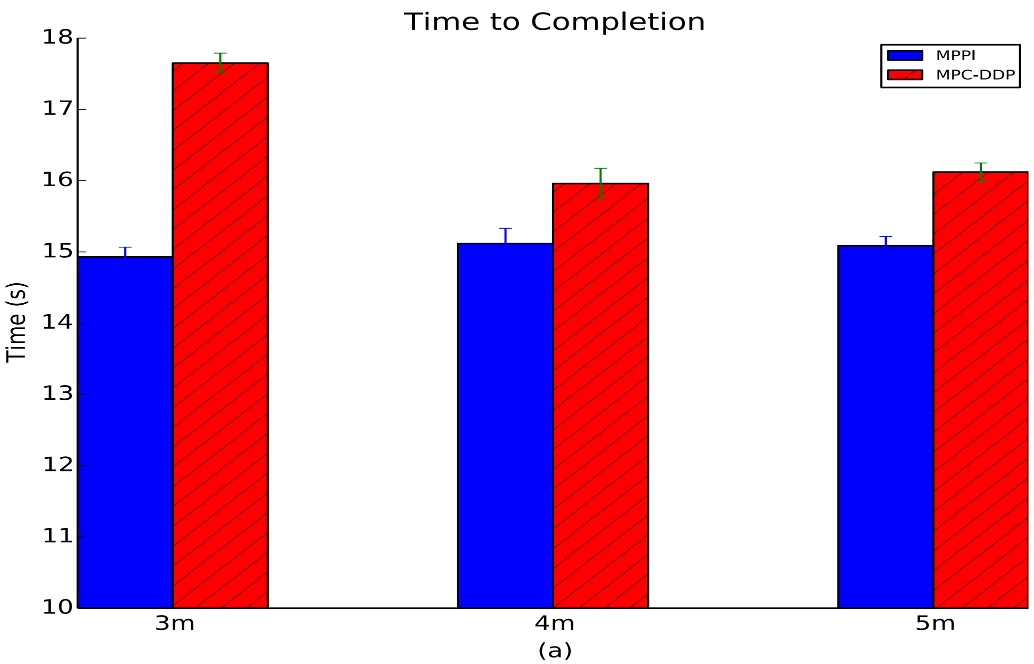
Applications in Multi-vehicle

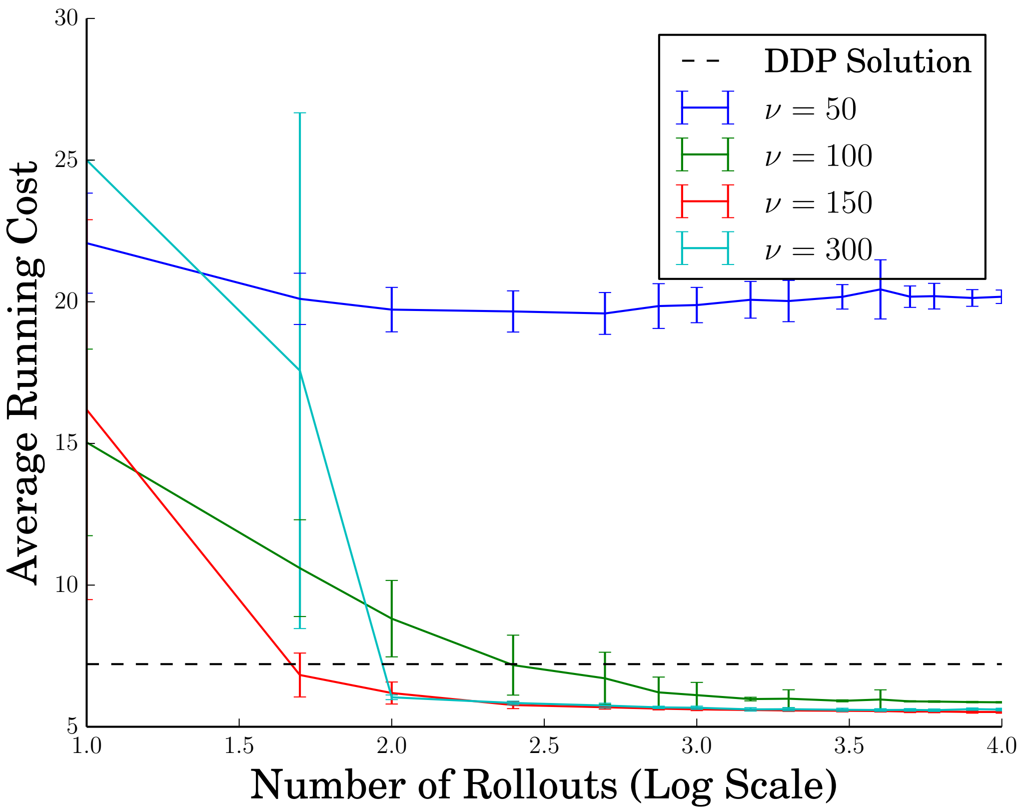


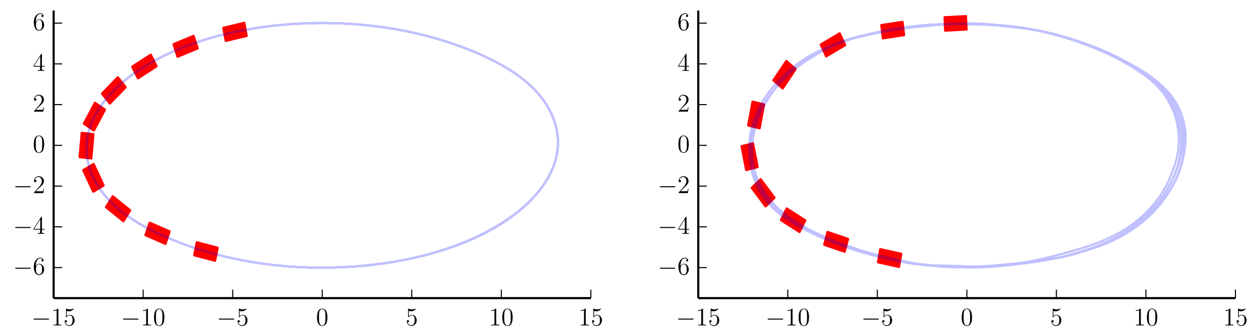
Applications in Robotics-Comparisons





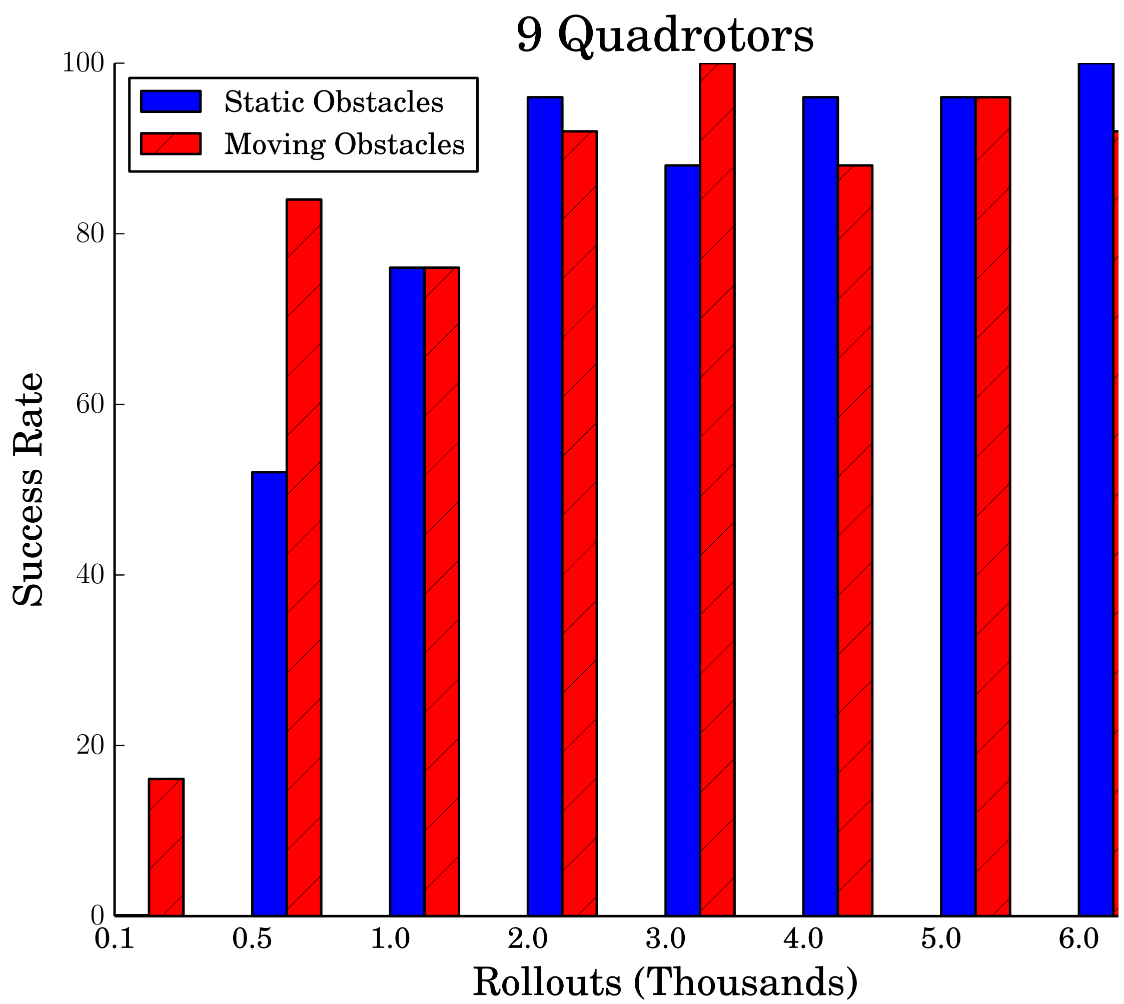


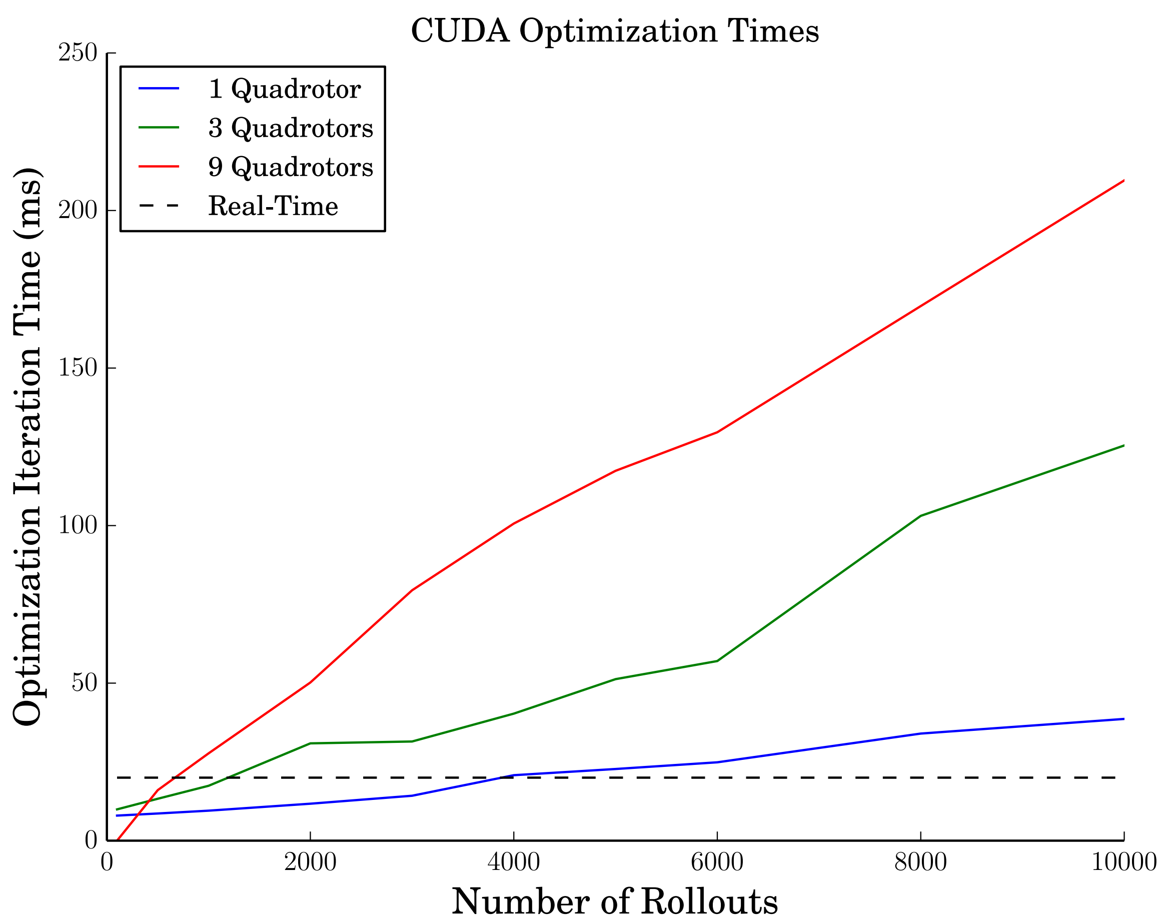




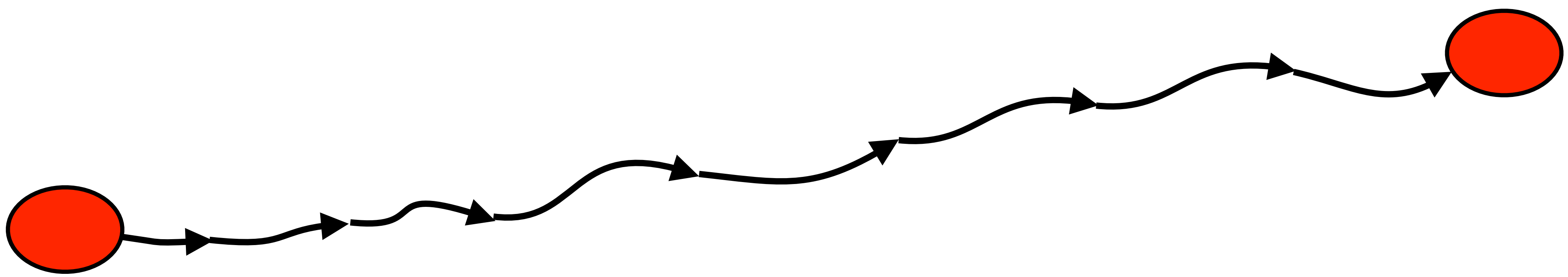
MPC-DDP **MPPI**

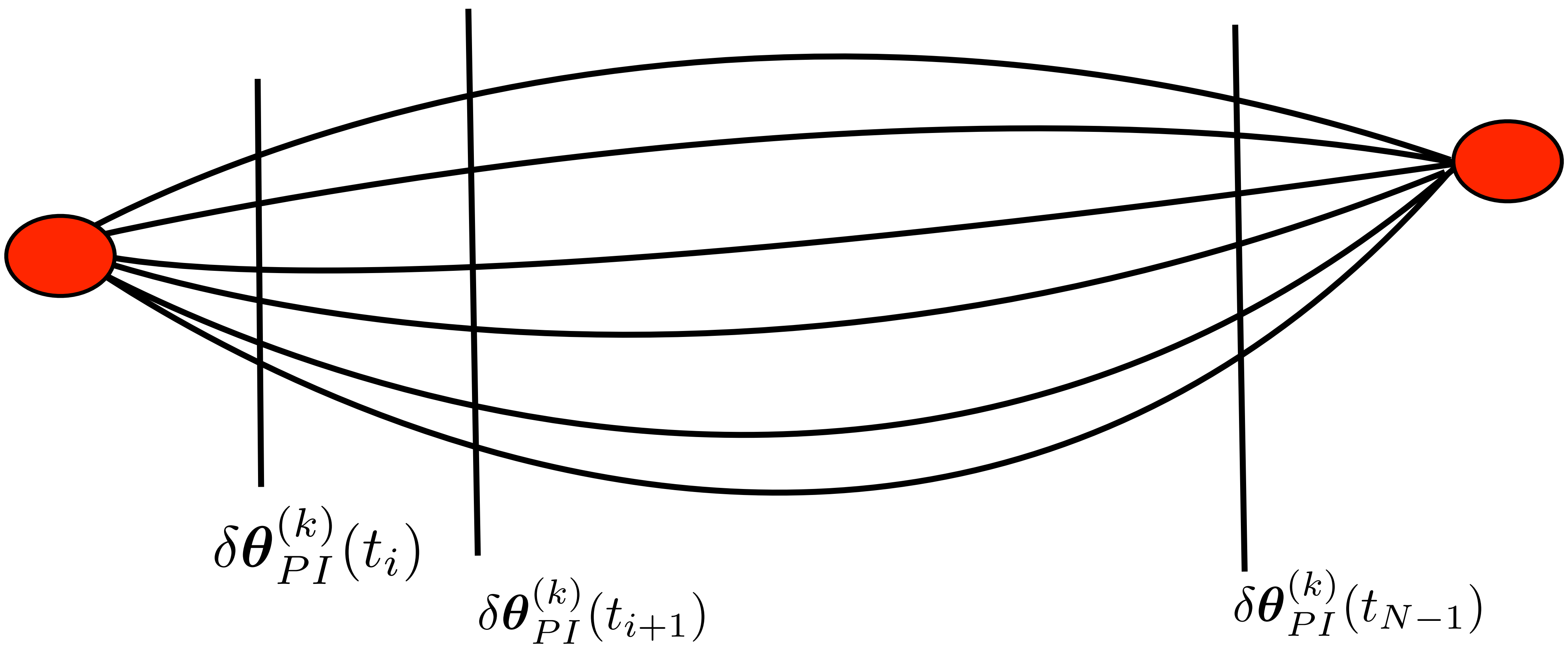
Applications in Multi-vehicle





Learning and optimization in different Time scales

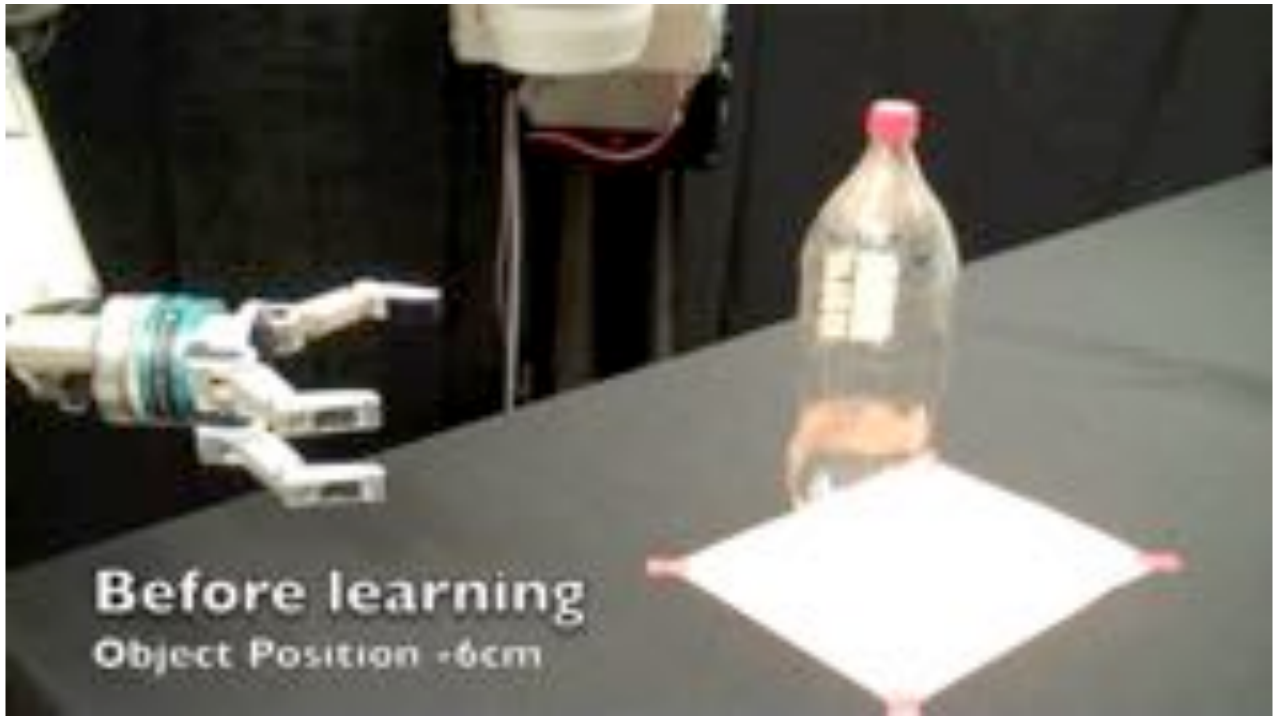




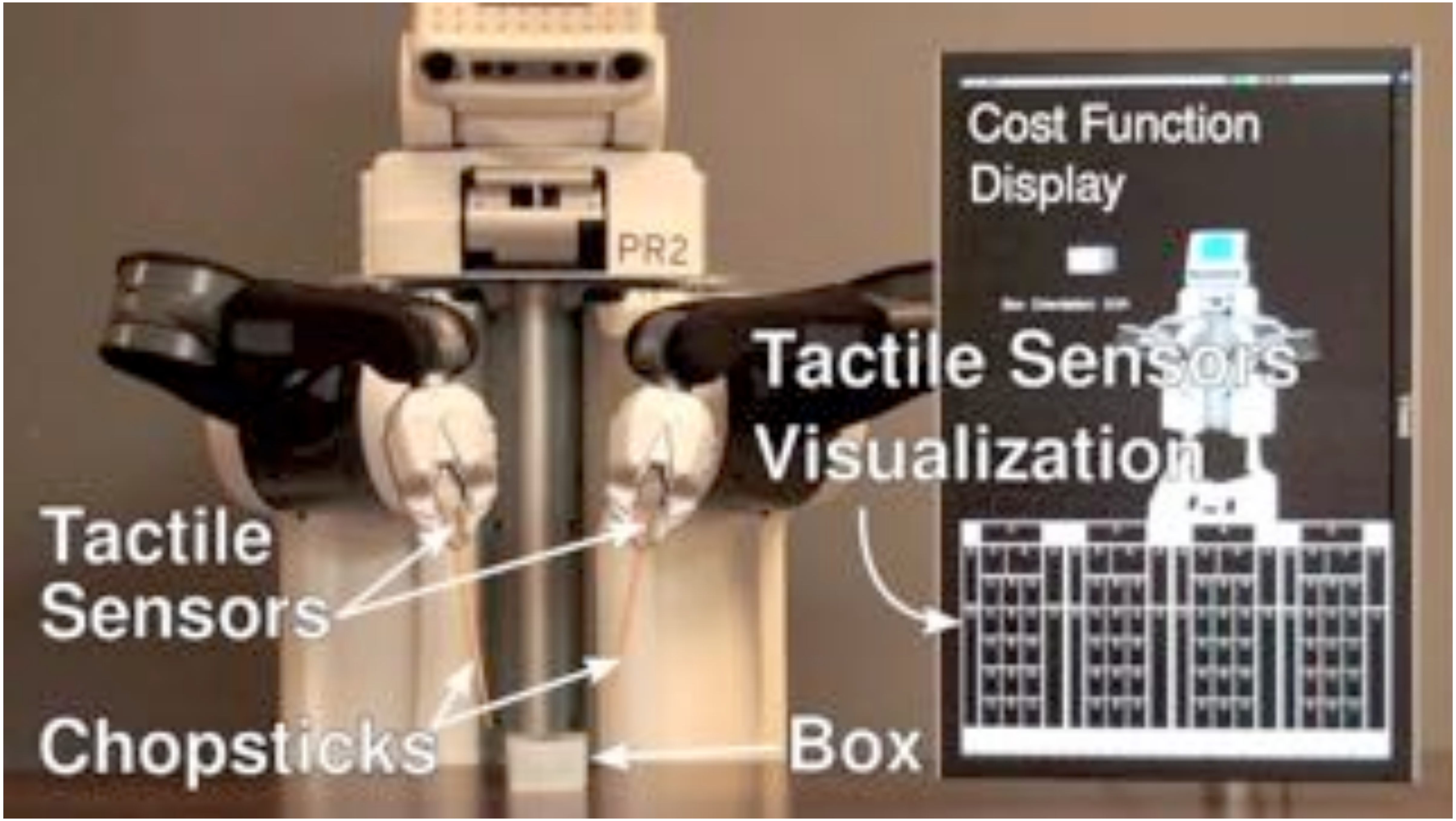
$$\mathbf{u}(\mathbf{x}, t, \boldsymbol{\theta}) = \boldsymbol{\Phi}(\mathbf{x}, t)\boldsymbol{\theta}$$

$$\mathbf{u}(\mathbf{x}, t, \theta) = \boldsymbol{\Phi}(\mathbf{x})\theta(\mathbf{t})$$

Applications in Robotics-Manipulation



Applications in Robotics-Manipulation



Nonlinear Feynman-Kac

Cost Function:

$$J(\tau, x_\tau; u(\cdot)) = \mathbb{E}\left[g(x(T)) + \int_\tau^T q_0(t, x(t)) + q_1(t, x(t))^\top |u(t)| \, \mathrm{d}t \right]$$

Stochastic Dynamics:

$$\mathrm{d}x(t) = f(t, x(t))\mathrm{d}t + G(t, x(t))u(t)\mathrm{d}t + \Sigma(t, x(t))\mathrm{d}W_t$$

Hamilton-Jacobi-Bellman:

$$v_t + \inf_{u \in U} \left\{ \frac{1}{2}\mathrm{tr}(v_{xx}\Sigma\Sigma^\top) + v_x^\top f + \big(v_x^\top G + q_1^\top \mathrm{D}(\mathrm{sgn}(u))\big)u + q_0 \right\} = 0$$

Hamilton-Jacobi-Bellman:

$$v_t + \frac{1}{2}\mathrm{tr}(v_{xx}\Sigma\Sigma^\top) + v_x^\top f + q_0 + \sum_{i=1}^{\nu} \min\left\{ (v_x^\top G + q_1^\top)_i u_i^{\max}, \ 0, \ -(v_x^\top G - q_1^\top)_i u_i^{\min} \right\} = 0$$

Nonlinear Feynman-Kac

Forward SDE: $\mathrm{d}X_s = b(s, X_s)\mathrm{d}s + \Sigma(s, X_s)\mathrm{d}W_s$

Backward SDE: $\mathrm{d}Y_s = -h(s, X_s^{t,x}, Y_s, Z_s)\mathrm{d}s + Z_s^\top \mathrm{d}W_s$

$$b(t,x) \equiv f(t,x)$$

$$h(t,x,z) \equiv q_0(t,x) + \sum_{i=1}^{\nu} \min\left\{ \left(z^\top \Gamma + q_1^\top\right)_i u_i^{\max},\ 0,\ \ -\left(z^\top \Gamma - q_1^\top\right)_i u_i^{\min} \right\}$$

Importance Forward SDE: $\mathrm{d}\tilde{X}_s = [b(s, \tilde{X}_s) + \Sigma(s, \tilde{X}_s)K_s]\mathrm{d}s + \Sigma(s, \tilde{X}_s)\mathrm{d}W_s$

Compensated Backward SDE: $\mathrm{d}\tilde{Y}_s = [-h(s, \tilde{X}_s, \tilde{Y}_s, \tilde{Z}_s) + \tilde{Z}_s^\top K_s]\mathrm{d}s + \tilde{Z}_s^\top \mathrm{d}W_s$

Nonlinear Feynman-Kac

Forward SDE: $\mathrm{d}X_s = b(s, X_s)\mathrm{d}s + \Sigma(s, X_s)\mathrm{d}W_s$

Backward SDE: $\mathrm{d}Y_s = -h(s, X_s^{t,x}, Y_s, Z_s)\mathrm{d}s + Z_s^\top \mathrm{d}W_s$

$$b(t,x) \equiv f(t,x)$$

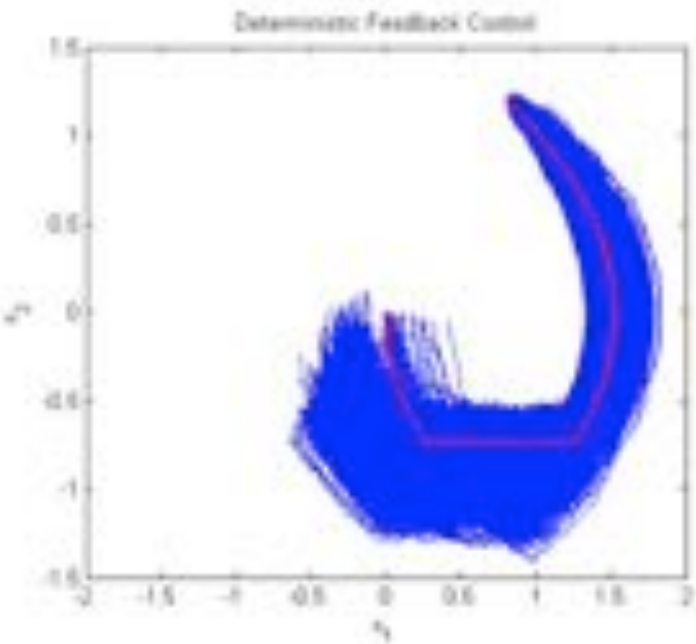
$$h(t,x,z) \equiv q_0(t,x) + \sum_{i=1}^{\nu} \min\left\{ \left(z^\top \Gamma + q_1^\top\right)_i u_i^{\max},\ 0,\ -\left(z^\top \Gamma - q_1^\top\right)_i u_i^{\min} \right\}$$

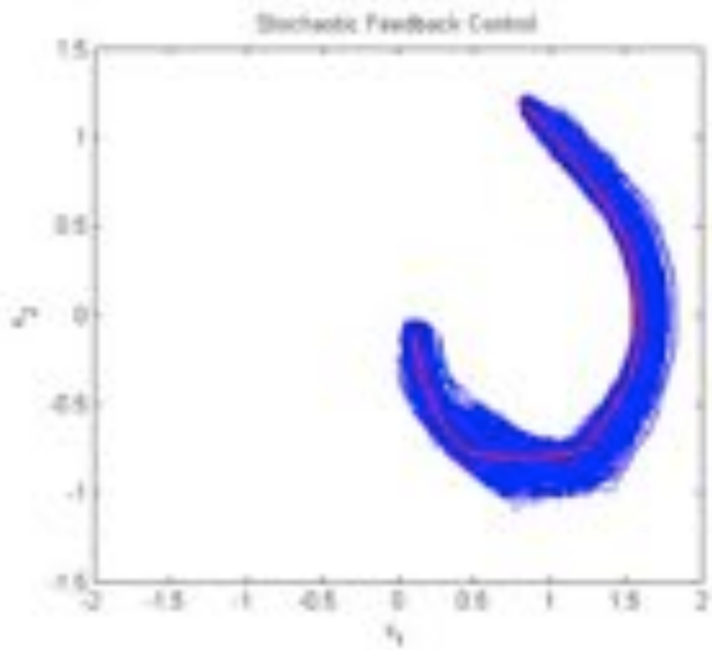
Importance Forward SDE: $\mathrm{d}\tilde{X}_s = [b(s, \tilde{X}_s) + \Sigma(s, \tilde{X}_s)K_s]\mathrm{d}s + \Sigma(s, \tilde{X}_s)\mathrm{d}W_s$

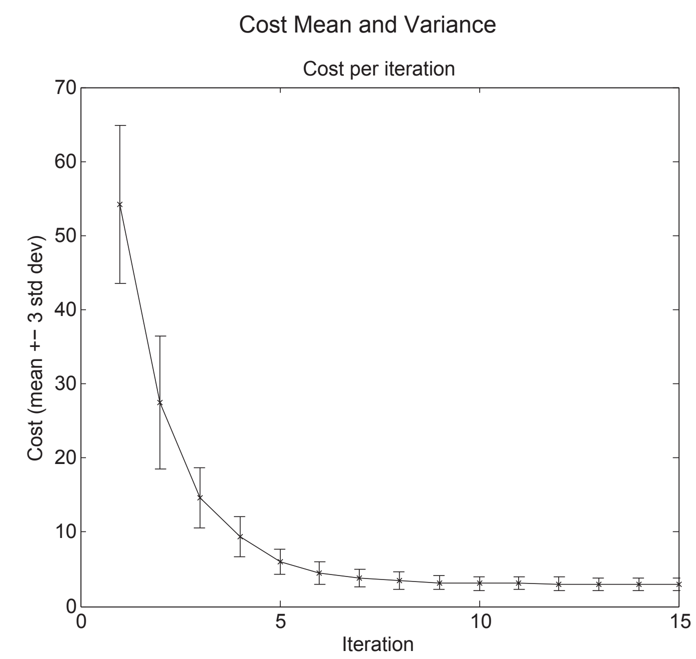
Compensated Backward SDE: $\mathrm{d}\tilde{Y}_s = [-h(s, \tilde{X}_s, \tilde{Y}_s, \tilde{Z}_s) + \tilde{Z}_s^\top K_s]\mathrm{d}s + \tilde{Z}_s^\top \mathrm{d}W_s$

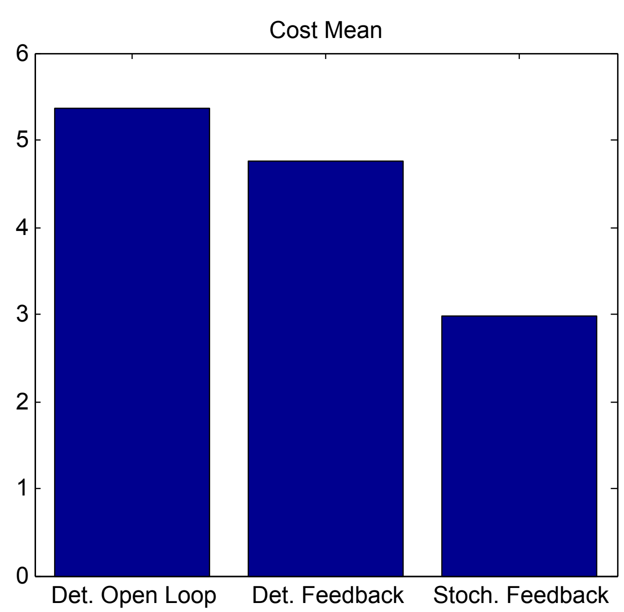
$$v_t + \frac{1}{2}\mathrm{tr}(v_{xx}\Sigma\Sigma^\top) + v_x^\top(b + \Sigma K) + h(t, x, v, \Sigma^\top v_x) - v_x^\top \Sigma K = 0$$

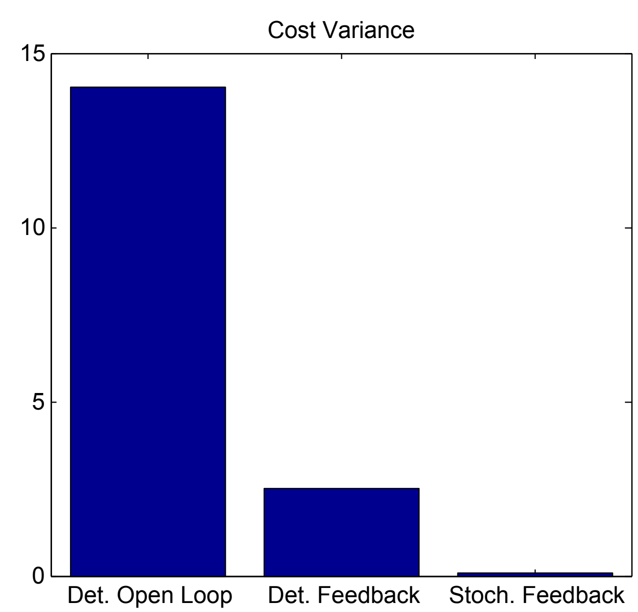
Nonlinear Feynman-Kac



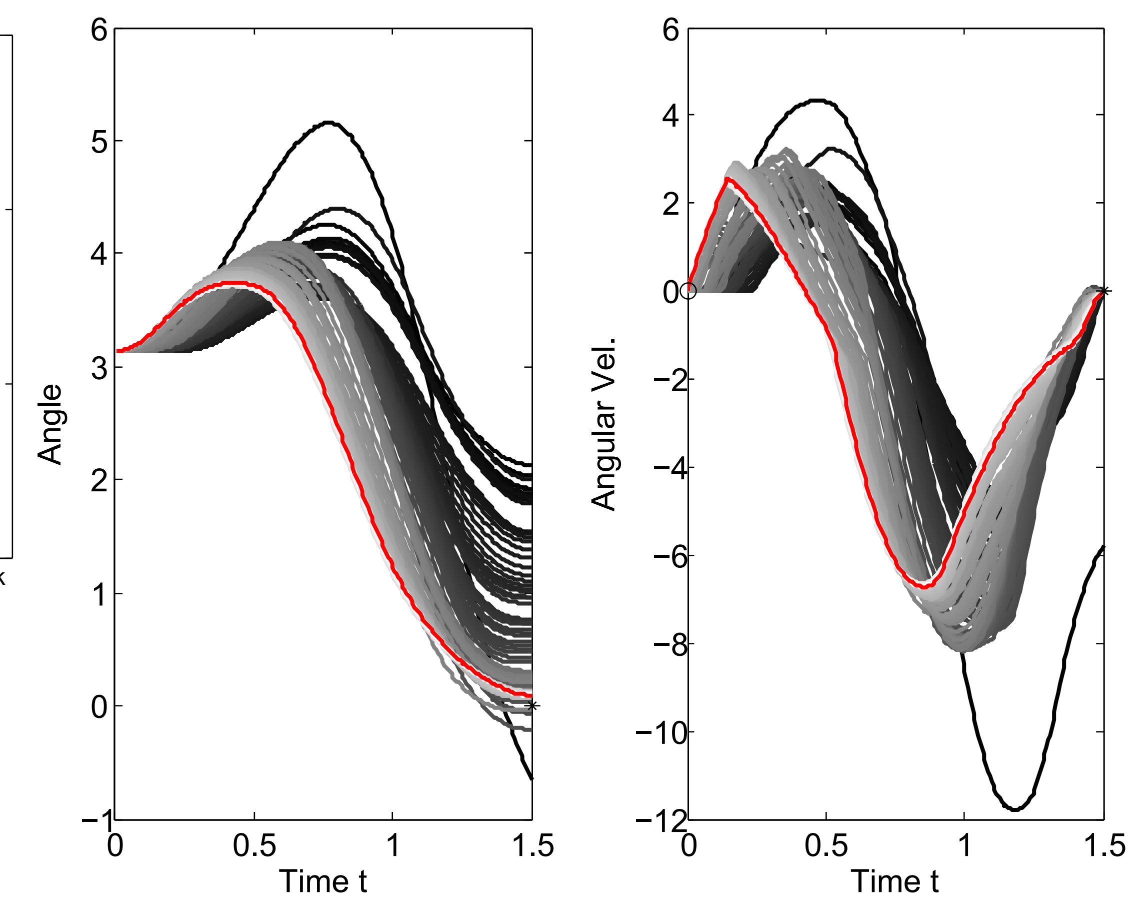








Mean of the Controlled System Trajectories of each Iteration

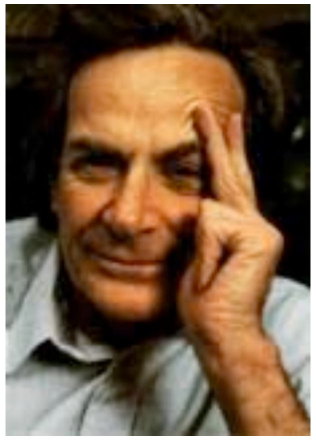


- **Stochastic Cooperative Games**
- **Risk Sensitive Stochastic**
- **Control affine, State/Control Multiplicative**

Georgia Institute of Technology

Summary

Statistical Physics



Richard Feynman

Mark Kac

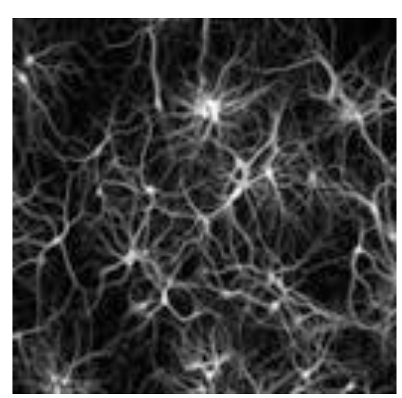
Control Theory

Richard Bellman

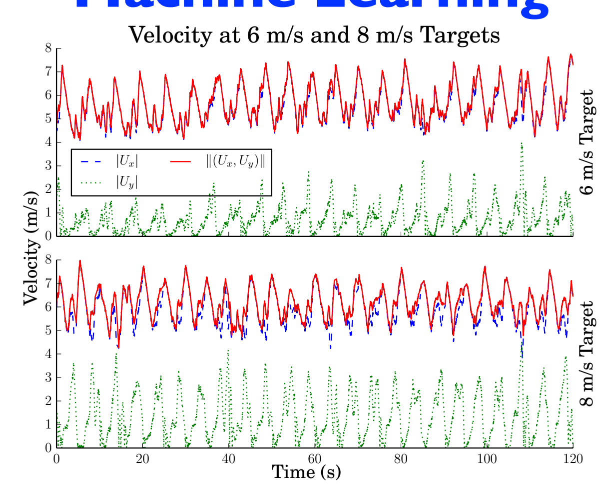
Lev Pontryagin

Parallel Computation

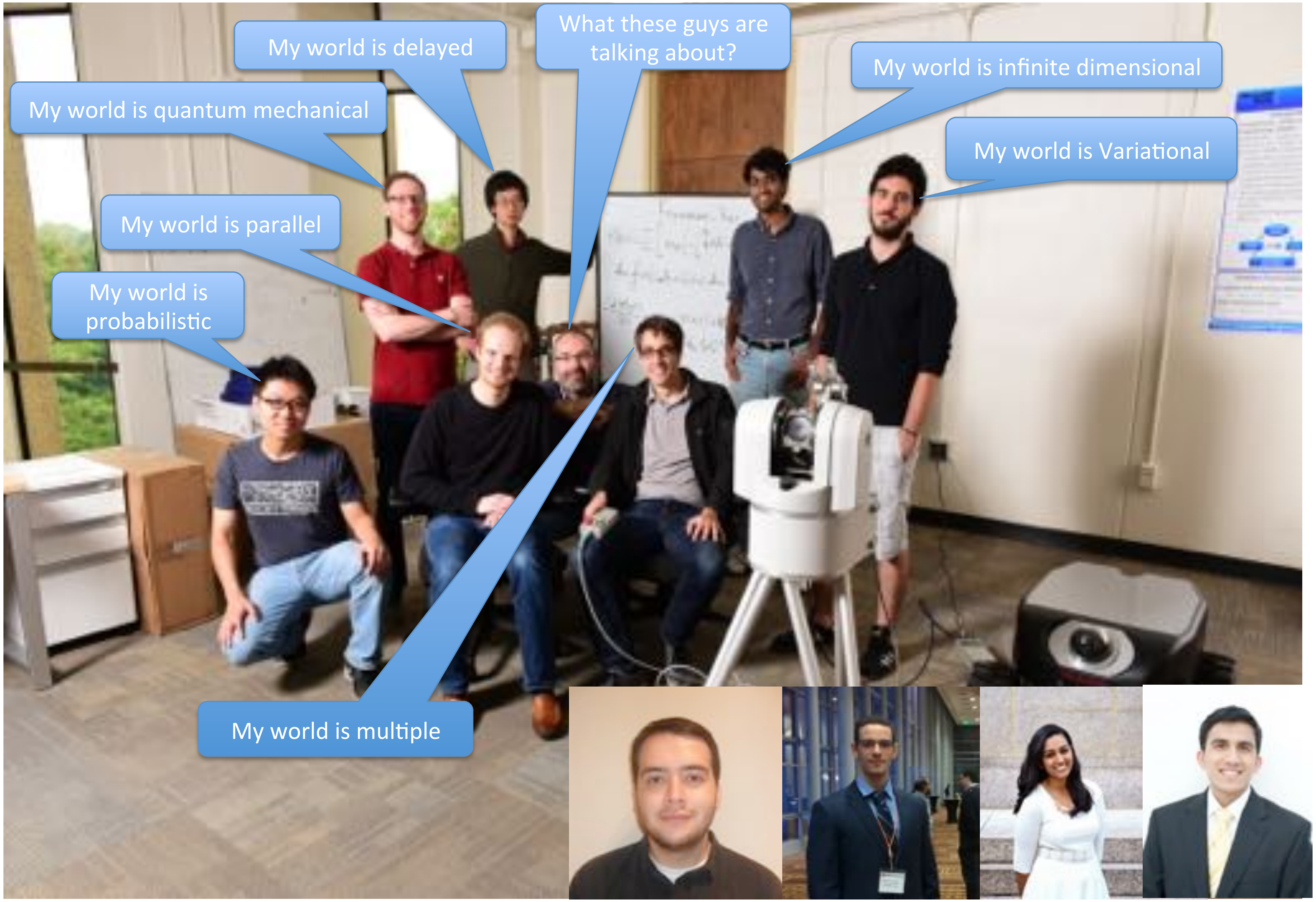
Neuro-Morphic



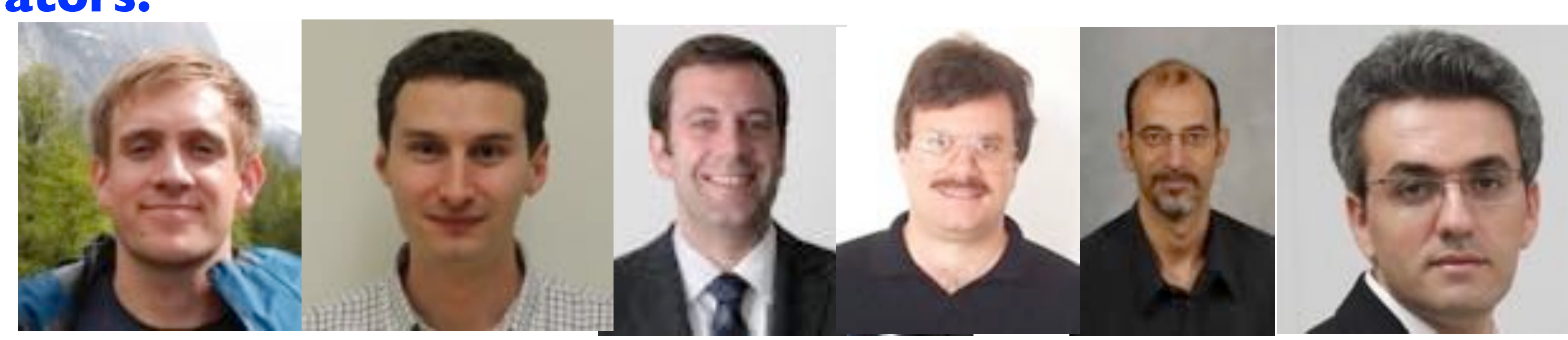
Machine Learning



Autonomous Control and Decision Systems Lab



Collaborators:



Sponsors:

Thanks !