Is Optimal Control a Panacea?

Evangelos Theodorou

University of Southern California

Now in University of Washington

Is optimal control a panacea?

Outline

Part I Model Based Nonlinear Optimal Control.

Constraint Iterative optimal control methods

Application on the neuromuscular model of the index finger

Conclusions.

Part II Path Integral Control Beyond Linearization and Internals Models.

Path Integral Stochastic Optimal Control

Application to a variety of dynamical systems

Conclusions.

Part III General comments about control under uncertainty.

Part I:
Model Based Nonlinear Optimal Control

Index Finger Biomechanics

Flexor Digitorum Superficialis(FDS)

Flexor Digitorum Profundus (FDP)

Extensor Communis(EC)

Extensor Indicis(EI)

Radial Interosseous(RI)

Ulnar Interosseous(UI)

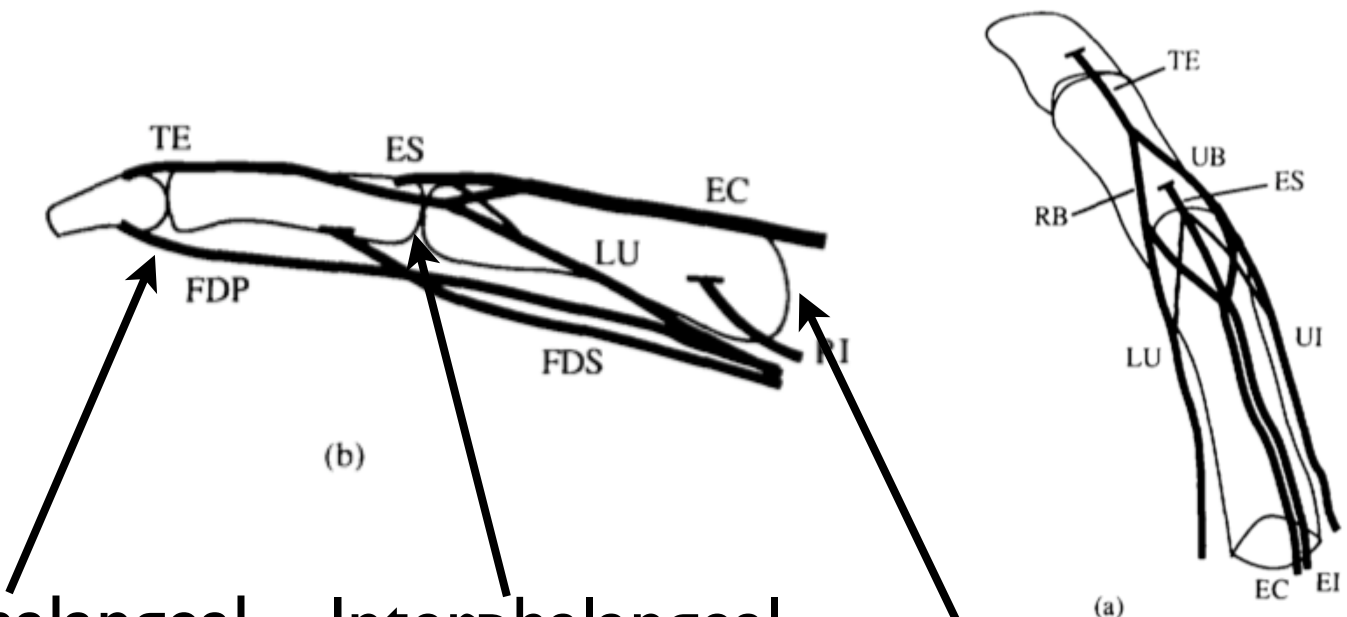
Lubrical(LU)

Terminal Extensor(TE)

Radial Band(RB)

Ulnar Band(UB)

Extensor Slip(ES)



Interphalangeal Distal

Interphalangeal Proximal

Metacarpophalangeal Proximal

N. Brook, J. Mizrahi, J. Dayan, 1995

Index Finger Model

$$\mathbf{T} = \mathbf{M}(\theta)\ \mathbf{F}$$

#torques x #tendons

$$\ddot{\boldsymbol{\theta}} = -\mathbf{I}\left(\boldsymbol{\theta}\right)^{-1}\left(\mathbf{C}\left(\boldsymbol{\theta},\dot{\boldsymbol{\theta}}\right) + \mathbf{B}\left(\boldsymbol{\theta}\right)\dot{\boldsymbol{\theta}} + \mathbf{T}\right)$$

$$\dot{\mathbf{F}} = -\frac{1}{\tau}\left(\mathbf{F} - \mathbf{u}\right)$$

$$J(\mathbf{x},\mathbf{u}) = \phi(\mathbf{x}(t_N)) + \int_0^{t_N} \mathbf{u}^T \mathbf{R} \mathbf{u} dt$$

E.A. Theodorou and FJ.Valero-Cuevas. EMBS2010

E.A. Theodorou, E. Todorov, and Francisco J.Valero-Cuevas. ASME 2010

E.A. Theodorou, Y. Tassa and E. Todorov. (ACC-2010).

E.A. Theodorou, E. Todorov, FJ Valero Cuevas. (ACC-2011)

Valero-Cuevas, F. J., Hoffmann, H., Kurse, M. U., Kutch, J. J., and Theodorou, E. A., 2009.. IEEE Reviews in Biomedical Engineering,

Nonlinear Stochastic Optimal Control

Given an initial control u(t), generate a trajectory x(t).

Start

Goal

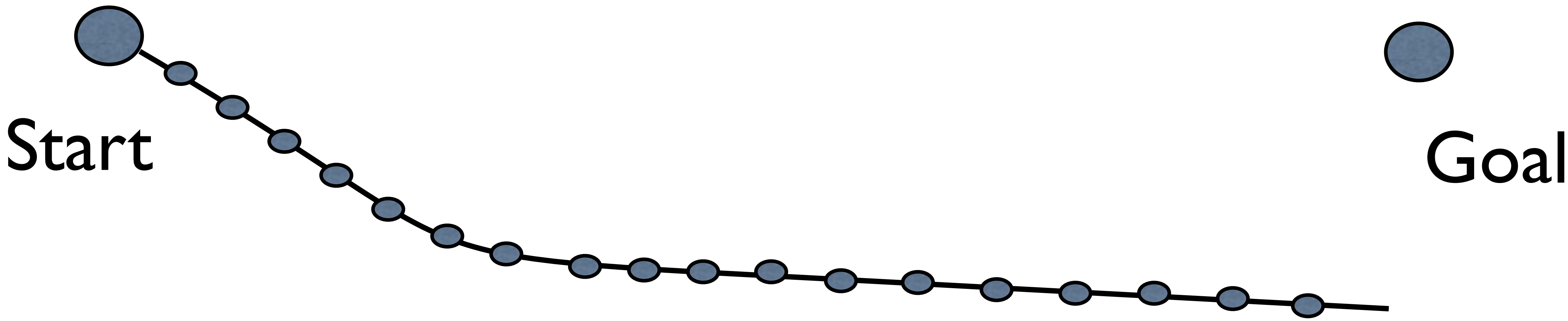
Step 1

E. Tododorov, W. Li ACC, ILQG, 2005.

D.H. Jacobson, D.Q. Mayne, Differential Dynamic Programming, 1970.

Nonlinear Stochastic Optimal Control

Given an initial control u(t), generate a trajectory x(t).



Step 1

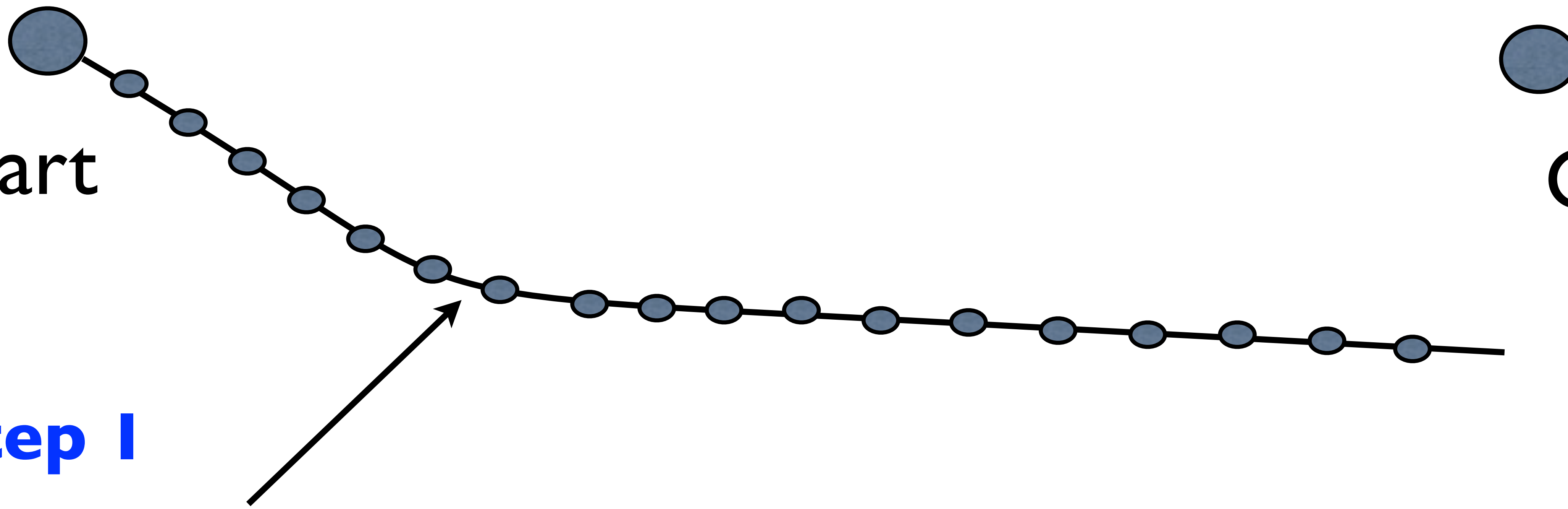
E. Tododorov, W. Li ACC, ILQG, 2005.

D.H. Jacobson, D.Q. Mayne, Differential Dynamic Programming, 1970.

Nonlinear Stochastic Optimal Control

Given an initial control u(t), generate a trajectory x(t).

Start



Goal

Step 1

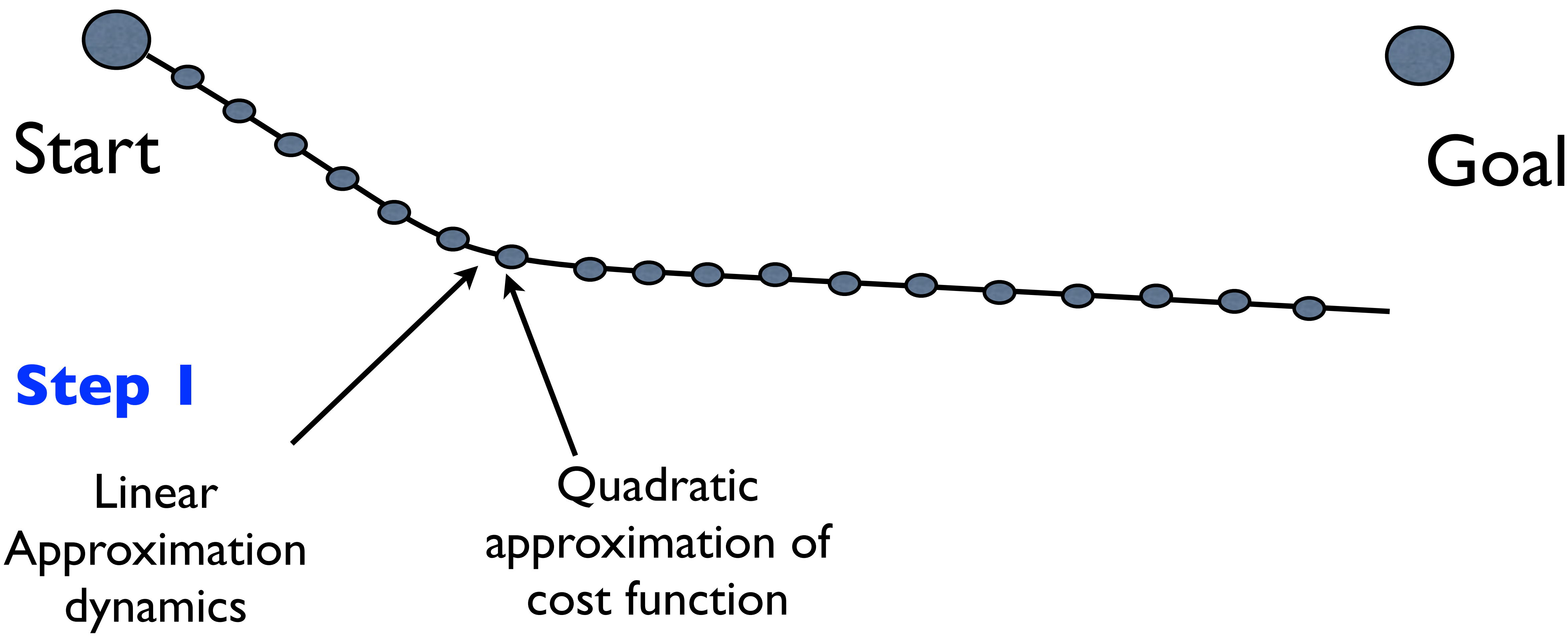
Linear
Approximation
dynamics

E. Tododorov, W. Li ACC, ILQG, 2005.

D.H. Jacobson, D.Q. Mayne, Differential Dynamic Programming, 1970.

Nonlinear Stochastic Optimal Control

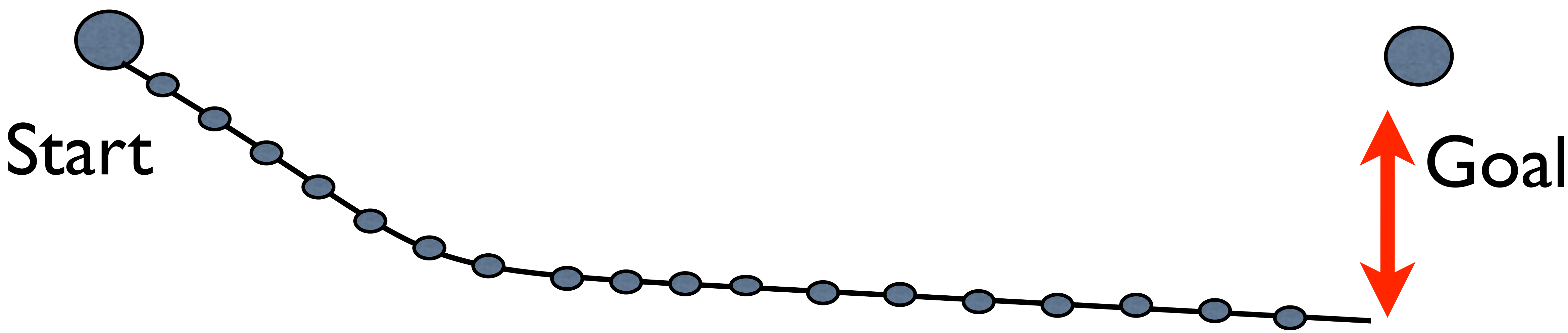
Given an initial control u(t), generate a trajectory x(t).



E. Tododorov, W. Li ACC, ILQG, 2005.

D.H. Jacobson, D.Q. Mayne, Differential Dynamic Programming, 1970.

Nonlinear Stochastic Optimal Control

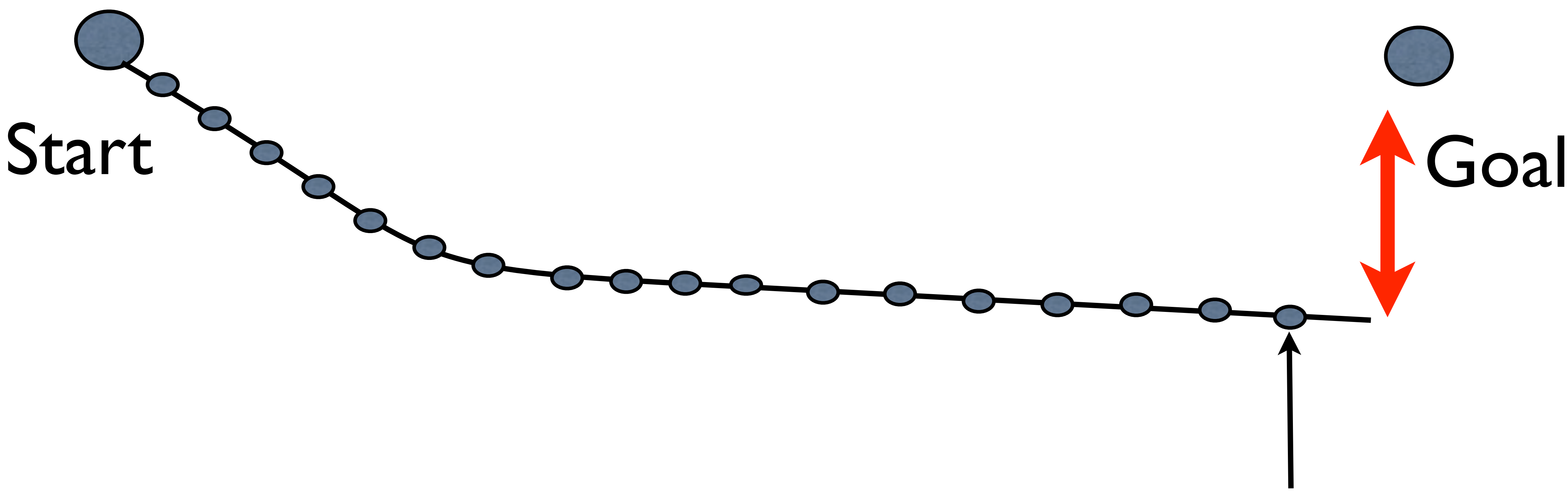


Step 2. Back-propagate the quadratic approximation of the Cost-to-go

E. Todorov, W. Li ACC, ILQG, 2005.

D.H. Jacobson, D.Q. Mayne, Differential Dynamic Programming, 1970.

Nonlinear Stochastic Optimal Control

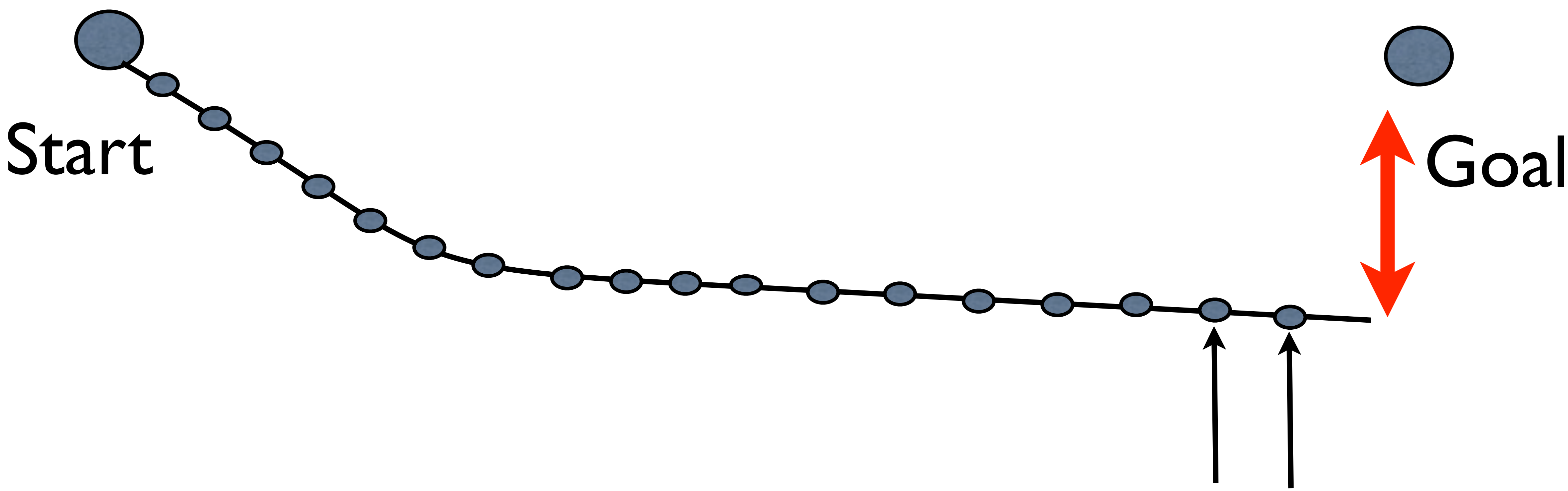


Step 2. Back-propagate the quadratic approximation of the Cost-to-go

E. Todorov, W. Li ACC, ILQG, 2005.

D.H. Jacobson, D.Q. Mayne, Differential Dynamic Programming, 1970.

Nonlinear Stochastic Optimal Control

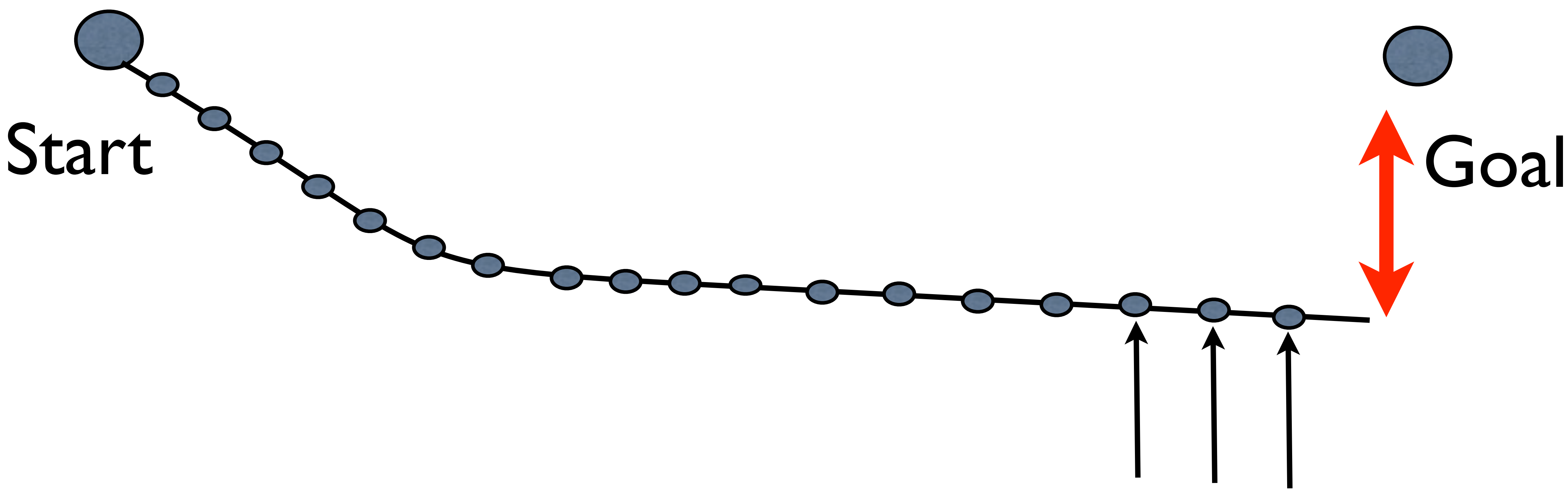


Step 2. Back-propagate the quadratic approximation of the Cost-to-go

E. Todorov, W. Li ACC, ILQG, 2005.

D.H. Jacobson, D.Q. Mayne, Differential Dynamic Programming, 1970.

Nonlinear Stochastic Optimal Control

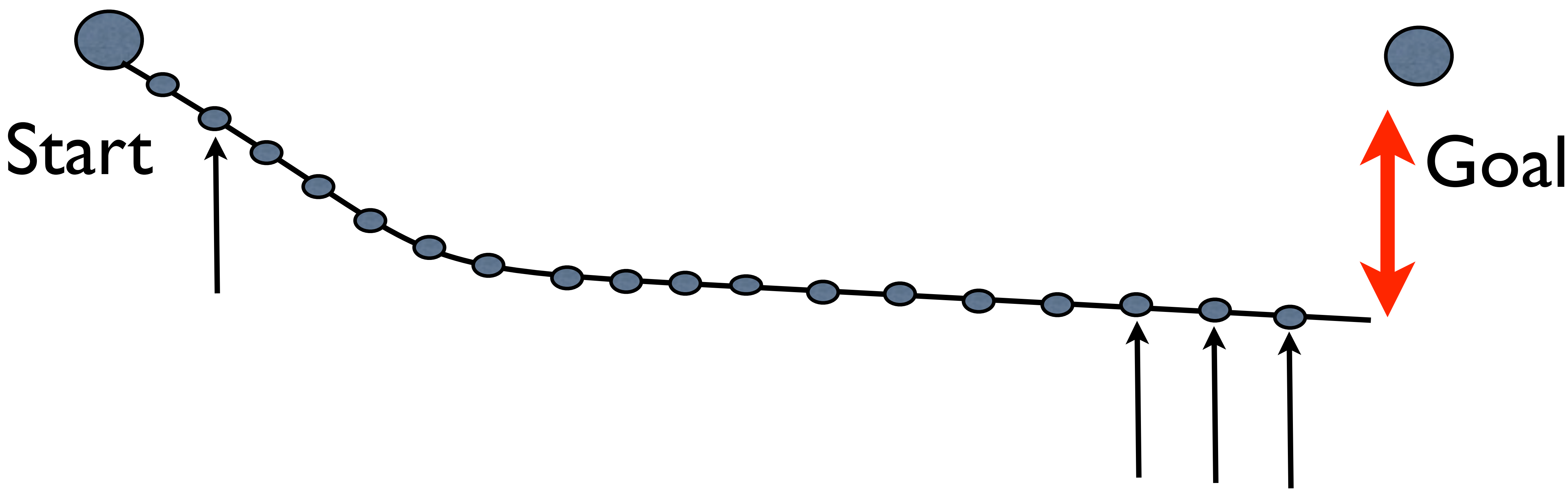


Step 2. Back-propagate the quadratic approximation of the Cost-to-go

E. Todorov, W. Li ACC, ILQG, 2005.

D.H. Jacobson, D.Q. Mayne, Differential Dynamic Programming, 1970.

Nonlinear Stochastic Optimal Control

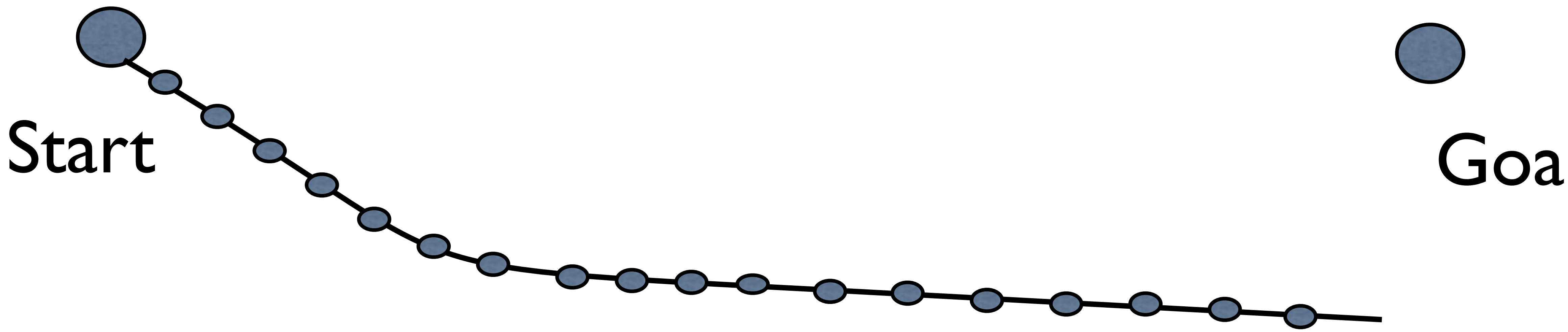


Step 2. Back-propagate the quadratic approximation of the Cost-to-go

E. Todorov, W. Li ACC, ILQG, 2005.

D.H. Jacobson, D.Q. Mayne, Differential Dynamic Programming, 1970.

Nonlinear Stochastic Optimal Control

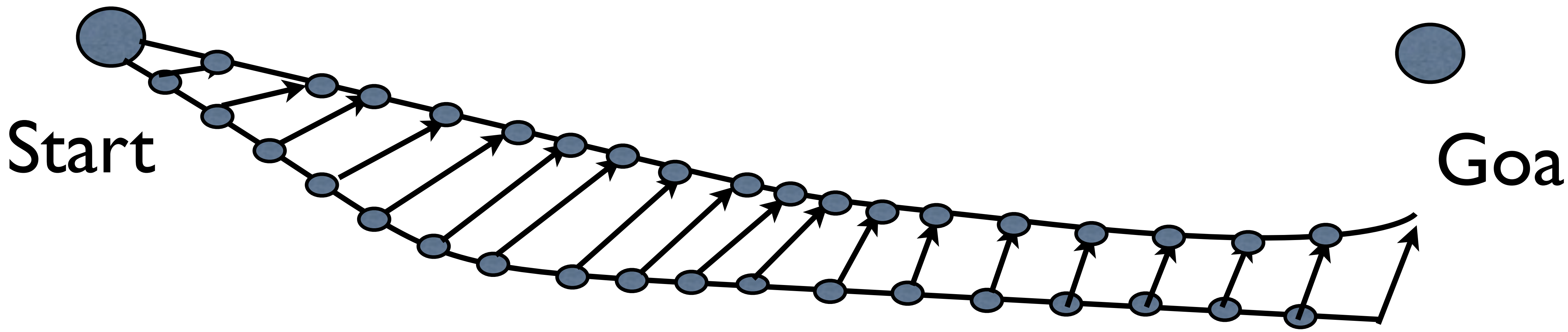


Step 3: Update the controls and the trajectory:

E. Todorov, W. Li ACC, ILQG, 2005.

D.H. Jacobson, D.Q. Mayne, Differential Dynamic Programming, 1970.

Nonlinear Stochastic Optimal Control

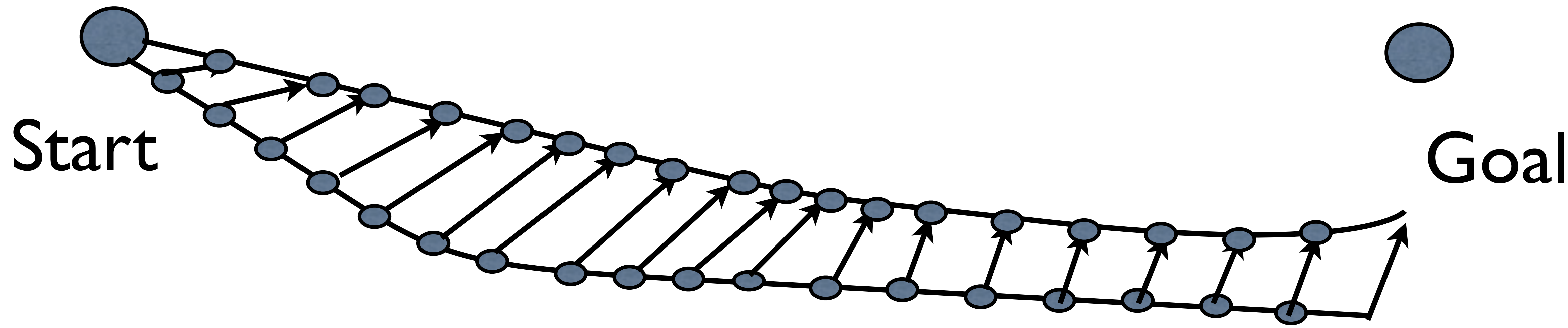


Step 3: Update the controls and the trajectory:

E. Todorov, W. Li ACC, ILQG, 2005.

D.H. Jacobson, D.Q. Mayne, Differential Dynamic Programming, 1970.

Nonlinear Stochastic Optimal Control



Step 3: Update the controls and the trajectory:

What you get at the end

Control gains L(t)

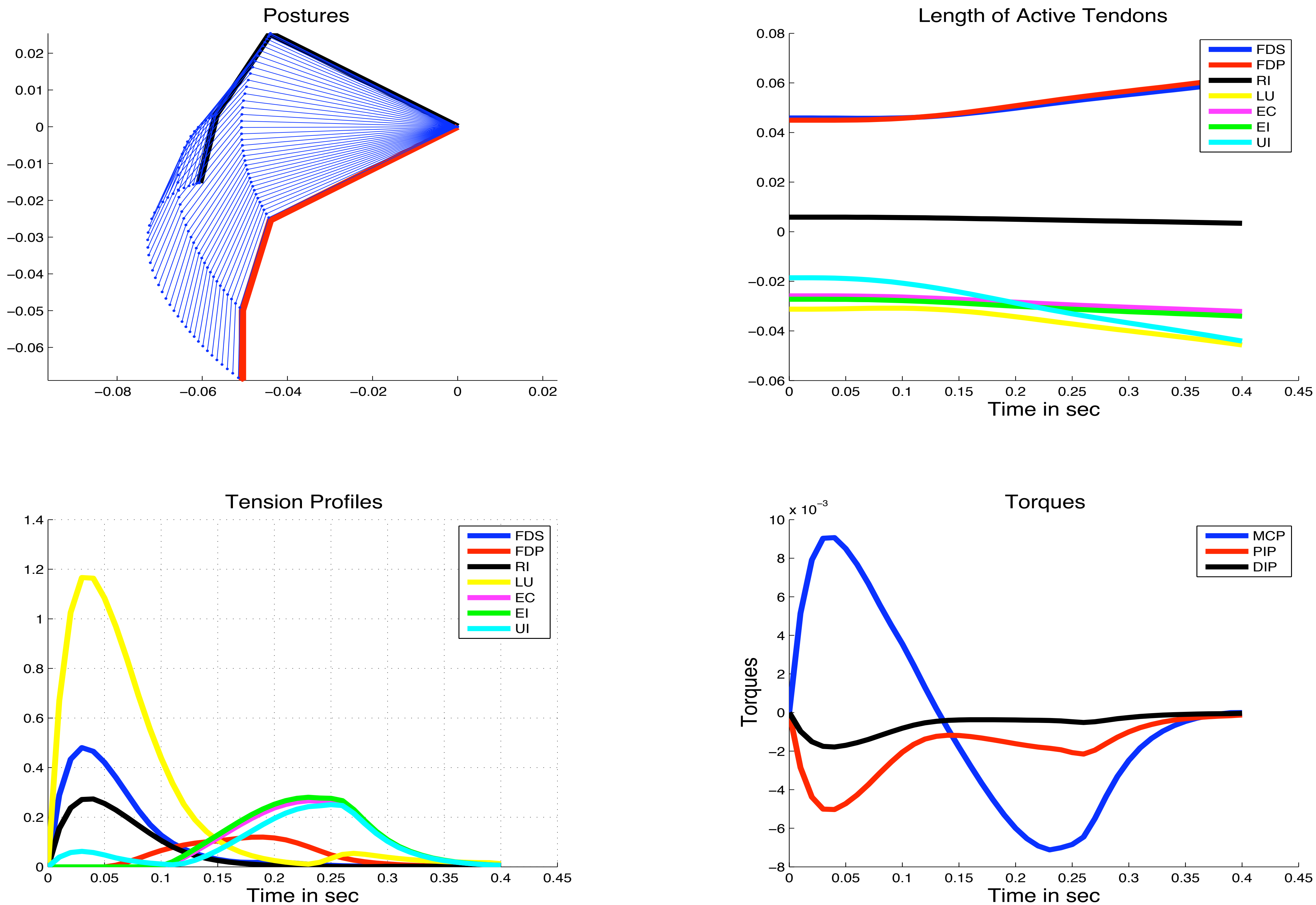
Optimal Trajectory x(t)

Optimal feedforward controls u(t)

E. Todorov, W. Li ACC, ILQG, 2005.

D.H. Jacobson, D.Q. Mayne, Differential Dynamic Programming, 1970.

Application on the index finger



Conclusions

Conclusions

Optimal control of tendon driven systems is a hard problem.

Conclusions

Optimal control of tendon driven systems is a hard problem.

Hard constrains in the actuations of the tendons reduce solution space.

Conclusions

Optimal control of tendon driven systems is a hard problem.

Hard constrains in the actuations of the tendons reduce solution space.

Complexity because tendons create torques around almost all joins.

Conclusions

Optimal control of tendon driven systems is a hard problem.

Hard constrains in the actuations of the tendons reduce solution space.

Complexity because tendons create torques around almost all joins.

There are many candidate moment arm models of the index finger.

Conclusions

Optimal control of tendon driven systems is a hard problem.

Hard constrains in the actuations of the tendons reduce solution space.

Complexity because tendons create torques around almost all joins.

There are many candidate moment arm models of the index finger.

Sensitivity of control policies to model changes and variations.

Conclusions

Optimal control of tendon driven systems is a hard problem.

Hard constrains in the actuations of the tendons reduce solution space.

Complexity because tendons create torques around almost all joins.

There are many candidate moment arm models of the index finger.

Sensitivity of control policies to model changes and variations.

Sensitivity increases due to linearization and hard constrains.

Conclusions

Optimal control of tendon driven systems is a hard problem.

Hard constrains in the actuations of the tendons reduce solution space.

Complexity because tendons create torques around almost all joins.

There are many candidate moment arm models of the index finger.

Sensitivity of control policies to model changes and variations.

Sensitivity increases due to linearization and hard constrains.

What you get at the end

Control gains L(t)

Optimal Trajectory x(t)

Optimal feedforward controls u(t)

Part 2:
Beyond linearization and internals model:
Path integral control approach

Path Integral Optimal Control

Dynamics: $$\mathbf{dx} = (\mathbf{f}(\mathbf{x}_t) + \mathbf{G}(\mathbf{x})\mathbf{u})\,dt + \mathbf{B}(\mathbf{x})\mathbf{L}d\omega$$

Cost: $$J(\mathbf{x},\mathbf{u}) = \phi(\mathbf{x}(t_N)) + \int_{t_0}^{t_N} q(\mathbf{x},t) + \mathbf{u}^T\mathbf{R}\mathbf{u}\;dt$$

Value Function: $$V(\mathbf{x},t) = \min_{\mathbf{u}} E\Big(J(\mathbf{x},\mathbf{u})\Big)$$

Optimal Control: $$\mathbf{u}(\mathbf{x},t) = -\mathbf{R}^{-1}\mathbf{G}^T grad\Big(V(\mathbf{x},t)\Big)$$

Path Integral Optimal Control

Dynamics: $\mathbf{dx} = (\mathbf{f}(\mathbf{x}_t) + \mathbf{G}(\mathbf{x})\mathbf{u})\,dt + \mathbf{B}(\mathbf{x})\mathbf{L}d\omega$

Cost: $J(\mathbf{x},\mathbf{u}) = \phi(\mathbf{x}(t_N)) + \int_{t_0}^{t_N} q(\mathbf{x},t) + \mathbf{u}^T\mathbf{R}\mathbf{u}\;dt$

Value Function: $V(\mathbf{x},t) = \min_{\mathbf{u}} E\Big(J(\mathbf{x},\mathbf{u})\Big)$

Optimal Control: $\mathbf{u}(\mathbf{x},t) = -\mathbf{R}^{-1}\mathbf{G}^T grad\Big(V(\mathbf{x},t)\Big)$

$$-\partial_t V_t = q_t + (\nabla_{\mathbf{x}} V_t)^T \mathbf{f}_t - \frac{1}{2}(\nabla_{\mathbf{x}} V_t)^T \mathbf{G}_t \mathbf{R}^{-1} \mathbf{G}_t^T (\nabla_{\mathbf{x}} V_t) + \frac{1}{2} trace\left((\nabla_{\mathbf{xx}} V_t)\mathbf{B}_t \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} \mathbf{B}_t^T\right)$$

Path Integral Optimal Control

Dynamics: $$\mathbf{dx} = (\mathbf{f}(\mathbf{x}_t) + \mathbf{G}(\mathbf{x})\mathbf{u})\,dt + \mathbf{B}(\mathbf{x})\mathbf{L}d\omega$$

Cost: $$J(\mathbf{x},\mathbf{u}) = \phi(\mathbf{x}(t_N)) + \int_{t_0}^{t_N} q(\mathbf{x},t) + \mathbf{u}^T\mathbf{R}\mathbf{u}\;dt$$

Value Function: $$V(\mathbf{x},t) = \min_{\mathbf{u}} E\Big(J(\mathbf{x},\mathbf{u})\Big)$$

Optimal Control: $$\mathbf{u}(\mathbf{x},t) = -\mathbf{R}^{-1}\mathbf{G}^T grad\Big(V(\mathbf{x},t)\Big)$$

$$-\partial_t V_t = q_t + (\nabla_{\mathbf{x}}V_t)^T\mathbf{f}_t - \underbrace{\boxed{\frac{1}{2}(\nabla_{\mathbf{x}}V_t)^T\mathbf{G}_t\mathbf{R}^{-1}\mathbf{G}_t^T(\nabla_{\mathbf{x}}V_t)}}_{\text{Nonlinear}} + \frac{1}{2}trace\left((\nabla_{\mathbf{xx}}V_t)\mathbf{B}_t\mathbf{\Sigma}_{\boldsymbol{\epsilon}}\mathbf{B}_t^T\right)$$

Path Integral Optimal Control

Dynamics: $$\mathbf{dx} = (\mathbf{f}(\mathbf{x}_t) + \mathbf{G}(\mathbf{x})\mathbf{u})\,dt + \mathbf{B}(\mathbf{x})\mathbf{L}d\omega$$

Cost: $$J(\mathbf{x},\mathbf{u}) = \phi(\mathbf{x}(t_N)) + \int_{t_0}^{t_N} q(\mathbf{x},t) + \mathbf{u}^T\mathbf{R}\mathbf{u}\;dt$$

Value Function: $$V(\mathbf{x},t) = \min_{\mathbf{u}} E\Big(J(\mathbf{x},\mathbf{u})\Big)$$

Optimal Control: $$\mathbf{u}(\mathbf{x},t) = -\mathbf{R}^{-1}\mathbf{G}^T grad\Big(V(\mathbf{x},t)\Big)$$

$$-\partial_t V_t = q_t + (\nabla_{\mathbf{x}} V_t)^T \mathbf{f}_t - \underbrace{\boxed{\frac{1}{2}(\nabla_{\mathbf{x}} V_t)^T \mathbf{G}_t \mathbf{R}^{-1}\mathbf{G}_t^T(\nabla_{\mathbf{x}} V_t)}}_{\text{Nonlinear}} + \underbrace{\boxed{\frac{1}{2} trace\left((\nabla_{\mathbf{xx}} V_t)\mathbf{B}_t \mathbf{\Sigma}_{\boldsymbol{\epsilon}} \mathbf{B}_t^T\right)}}_{\text{Second order}}$$

Path Integral Optimal Control

Dynamics: $\mathbf{dx} = \left(\mathbf{f}(\mathbf{x}_t) + \mathbf{G}(\mathbf{x})\mathbf{u}\right) dt + \mathbf{B}(\mathbf{x})\mathbf{L} d\omega$

Cost: $J(\mathbf{x}, \mathbf{u}) = \phi(\mathbf{x}(t_N)) + \int_{t_0}^{t_N} q(\mathbf{x}, t) + \mathbf{u}^T \mathbf{R} \mathbf{u} \; dt$

Value Function: $V(\mathbf{x}, t) = \min_{\mathbf{u}} E \Big(J(\mathbf{x}, \mathbf{u}) \Big)$

Optimal Control: $\mathbf{u}(\mathbf{x}, t) = -\mathbf{R}^{-1} \mathbf{G}^T grad \Big(V(\mathbf{x}, t) \Big)$

$$-\partial_t V_t = q_t + (\nabla_{\mathbf{x}} V_t)^T \mathbf{f}_t - \underbrace{\boxed{\frac{1}{2} (\nabla_{\mathbf{x}} V_t)^T \mathbf{G}_t \mathbf{R}^{-1} \mathbf{G}_t^T (\nabla_{\mathbf{x}} V_t)}}_{\text{Nonlinear}} + \underbrace{\boxed{\frac{1}{2} trace \left((\nabla_{\mathbf{xx}} V_t) \mathbf{B}_t \mathbf{\Sigma}_{\boldsymbol{\epsilon}} \mathbf{B}_t^T \right)}}_{\text{Second order}}$$

We can simplify the problem by transforming HJB into a linear PDE.

Path Integral Optimal Control

Dynamics: $$\mathbf{dx} = (\mathbf{f}(\mathbf{x}_t) + \mathbf{G}(\mathbf{x})\mathbf{u})\, dt + \mathbf{B}(\mathbf{x})\mathbf{L}d\omega$$

Cost: $$J(\mathbf{x},\mathbf{u}) = \phi(\mathbf{x}(t_N)) + \int_{t_0}^{t_N} q(\mathbf{x},t) + \mathbf{u}^T\mathbf{R}\mathbf{u}\; dt$$

Value Function: $$V(\mathbf{x},t) = \min_{\mathbf{u}} E\Big(J(\mathbf{x},\mathbf{u})\Big)$$

Optimal Control: $$\mathbf{u}(\mathbf{x},t) = -\mathbf{R}^{-1}\mathbf{G}^T grad\Big(V(\mathbf{x},t)\Big)$$

$$-\partial_t V_t = q_t + (\nabla_{\mathbf{x}} V_t)^T \mathbf{f}_t - \underbrace{\boxed{\frac{1}{2}(\nabla_{\mathbf{x}} V_t)^T \mathbf{G}_t \mathbf{R}^{-1}\mathbf{G}_t^T (\nabla_{\mathbf{x}} V_t)}}_{\text{Nonlinear}} + \underbrace{\boxed{\frac{1}{2} trace\left((\nabla_{\mathbf{xx}} V_t)\mathbf{B}_t \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} \mathbf{B}_t^T\right)}}_{\text{Second order}}$$

We can simplify the problem by transforming HJB into a linear PDE.

We derive the path integral optimal controller.

Path Integral Optimal Control

Dynamics: $$d\mathbf{x} = (\mathbf{f}(\mathbf{x}_t) + \mathbf{G}(\mathbf{x})\mathbf{u})\, dt + \mathbf{B}(\mathbf{x})\mathbf{L} d\omega$$

Cost: $$J(\mathbf{x},\mathbf{u}) = \phi(\mathbf{x}(t_N)) + \int_{t_0}^{t_N} q(\mathbf{x},t) + \mathbf{u}^T\mathbf{R}\mathbf{u}\; dt$$

Value Function: $$V(\mathbf{x},t) = \min_{\mathbf{u}} E\Big(J(\mathbf{x},\mathbf{u}) \Big)$$

Optimal Control: $$\mathbf{u}(\mathbf{x},t) = -\mathbf{R}^{-1}\mathbf{G}^T grad\Big(V(\mathbf{x},t) \Big)$$

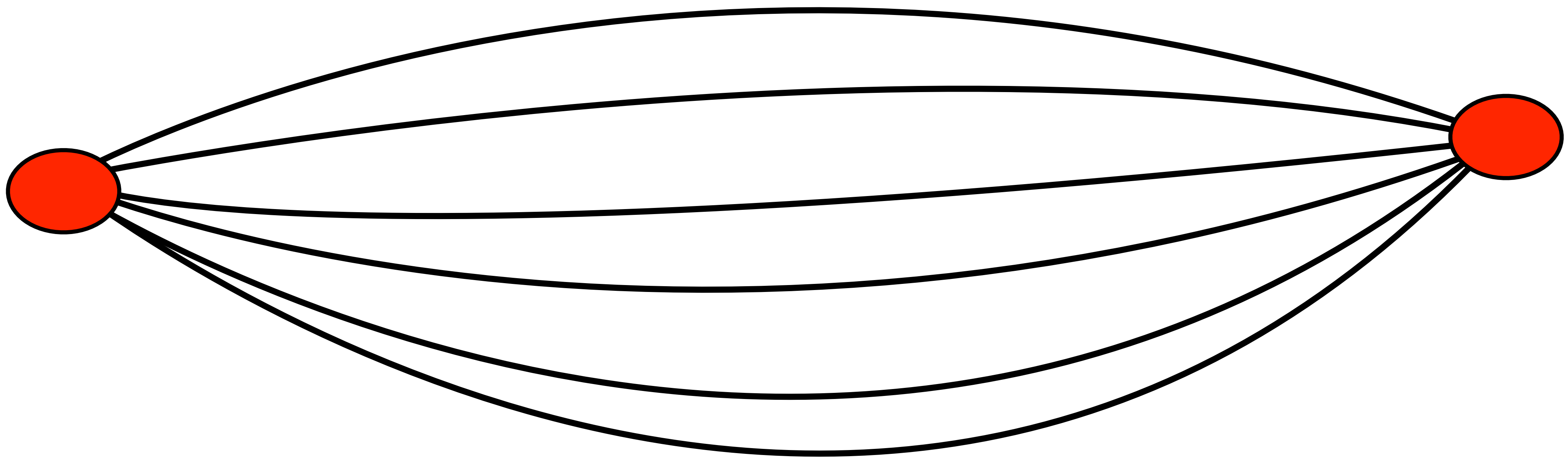
$$-\partial_t V_t = q_t + (\nabla_{\mathbf{x}} V_t)^T \mathbf{f}_t - \underbrace{\boxed{\frac{1}{2}(\nabla_{\mathbf{x}} V_t)^T \mathbf{G}_t \mathbf{R}^{-1} \mathbf{G}_t^T (\nabla_{\mathbf{x}} V_t)}}_{\text{Nonlinear}} + \underbrace{\boxed{\frac{1}{2} trace\left((\nabla_{\mathbf{xx}} V_t)\mathbf{B}_t \mathbf{\Sigma}_{\boldsymbol{\epsilon}} \mathbf{B}_t^T\right)}}_{\text{Second order}}$$

We can simplify the problem by transforming HJB into a linear PDE.

We derive the path integral optimal controller.

Path integral optimal control can be model based, semi-model based, model free.

Path Integral Optimal Control

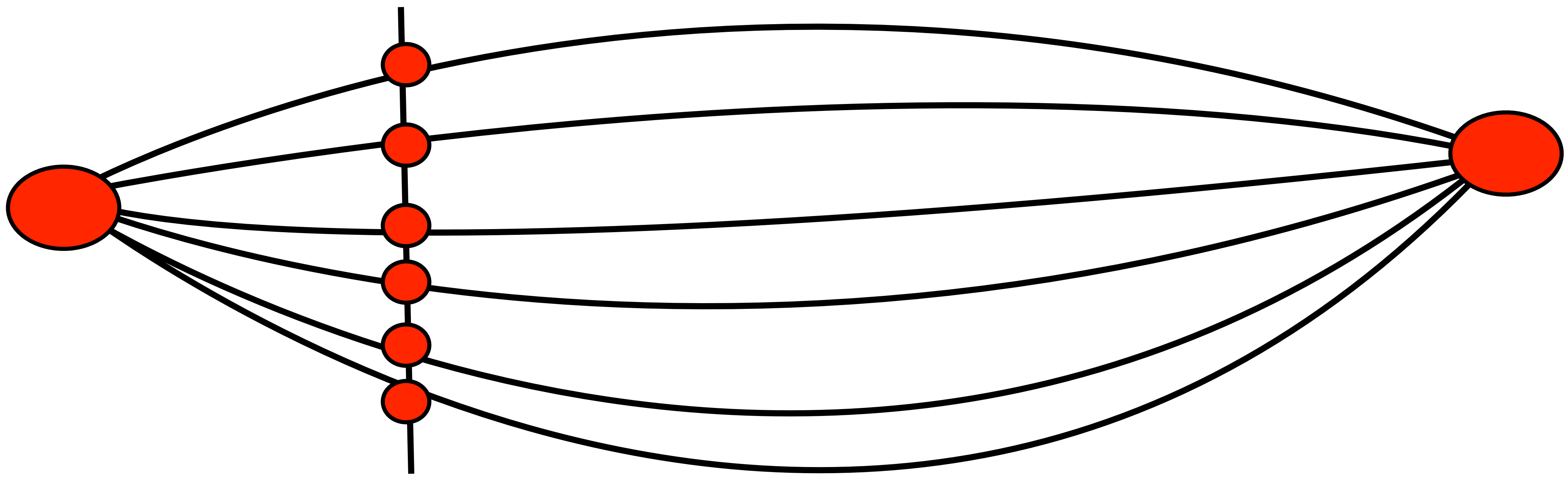


E. Theodorou, J. Buchli, S.Schaal ICRA 2010

E. Theodorou, J. Buchli, S.Schaal AISTATS 2010

E. Theodorou, J. Buchli, S.Schaal JMLR 2010

Path Integral Optimal Control

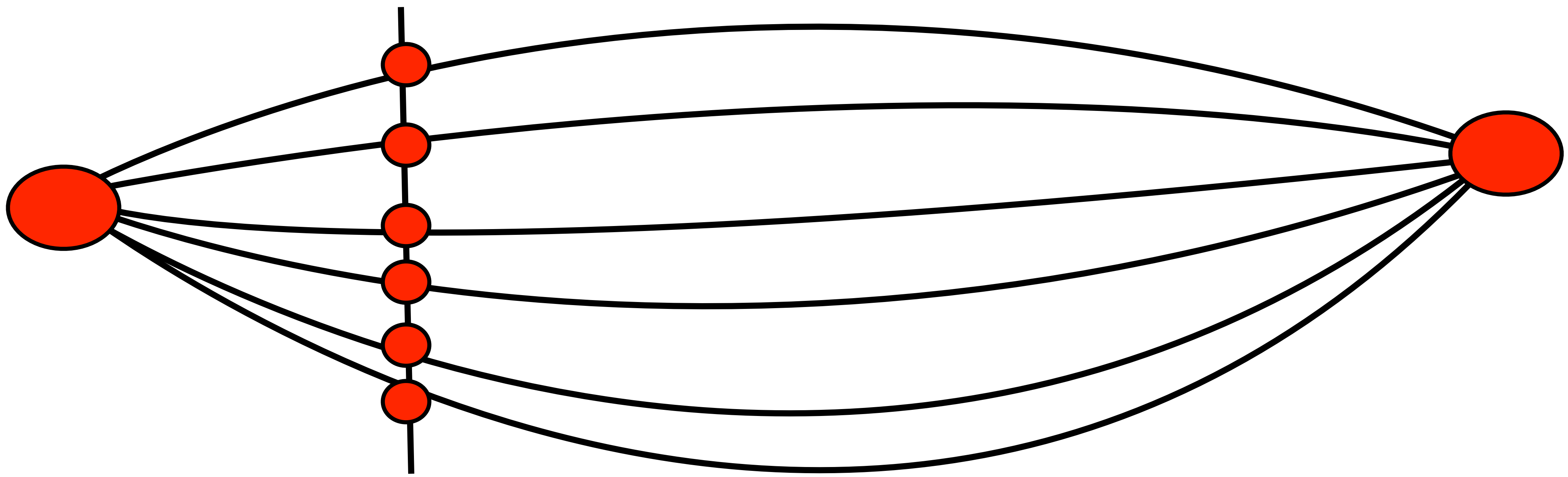


E. Theodorou, J. Buchli, S.Schaal ICRA 2010

E. Theodorou, J. Buchli, S.Schaal AISTATS 2010

E. Theodorou, J. Buchli, S.Schaal JMLR 2010

Path Integral Optimal Control



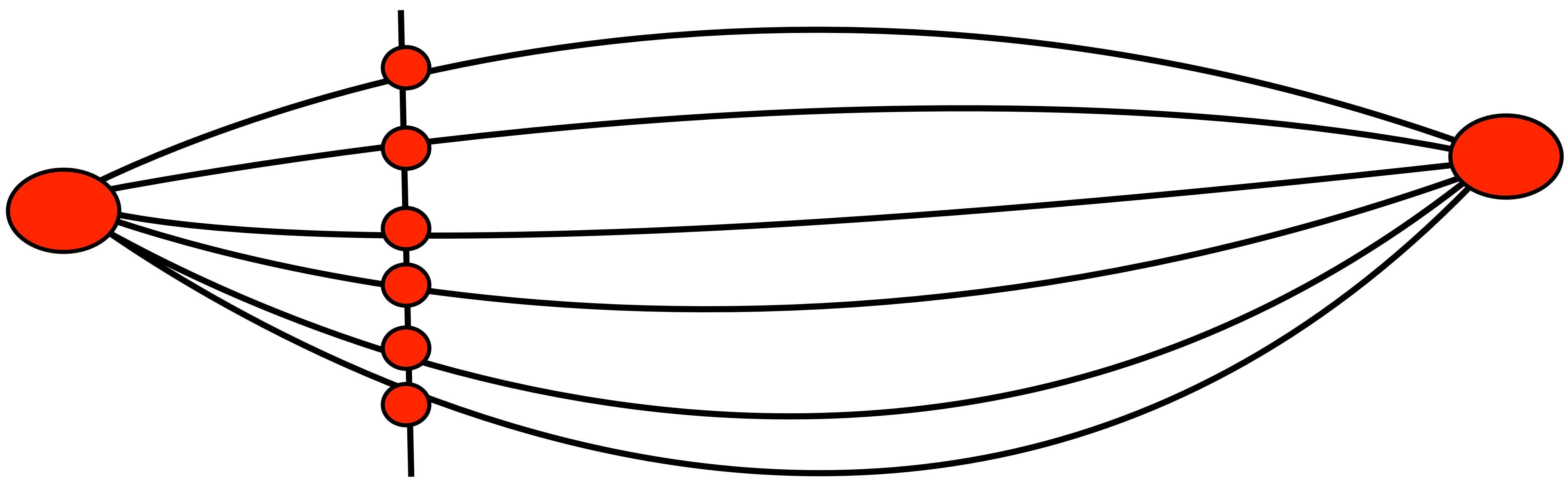
Cost $J(\mathbf{x}_j(t:N))$

E. Theodorou, J. Buchli, S.Schaal ICRA 2010

E. Theodorou, J. Buchli, S.Schaal AISTATS 2010

E. Theodorou, J. Buchli, S.Schaal JMLR 2010

Path Integral Optimal Control



Cost $J(\mathbf{x}_j(t:N))$

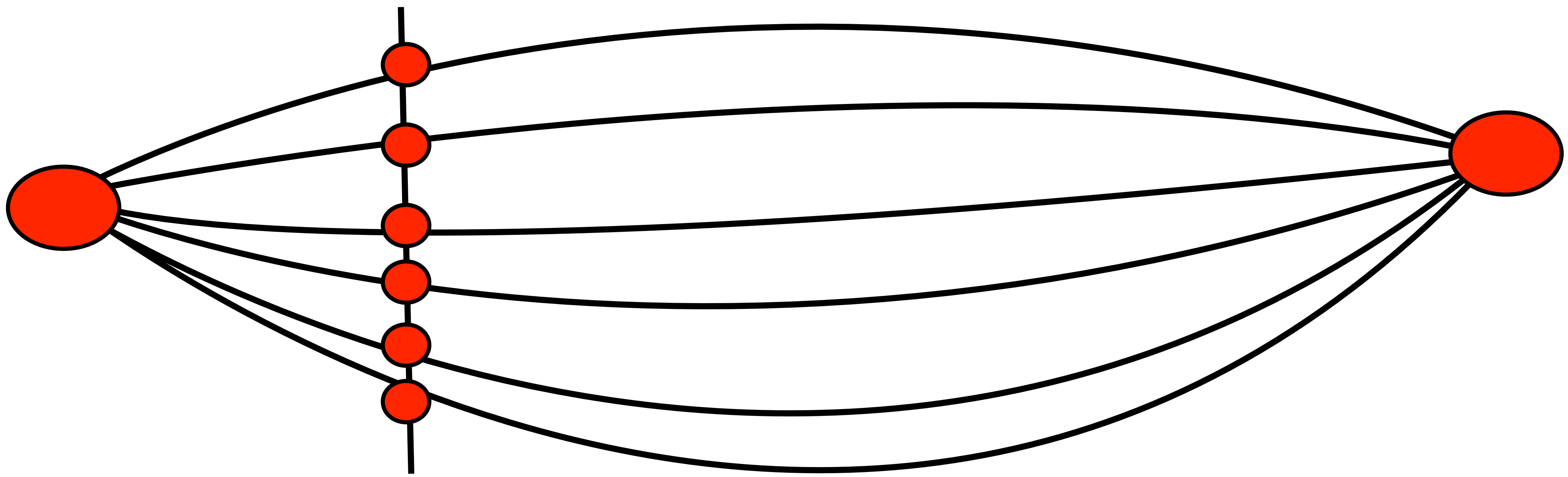
Probability $P(\mathbf{x}_j(t:N)) = \frac{e^{-\frac{1}{\lambda}P(\mathbf{x}_j(t:N))}}{\int e^{-\frac{1}{\lambda}P(\mathbf{x}_j(t:N))}d\mathbf{x}}$

E. Theodorou, J. Buchli, S.Schaal ICRA 2010

E. Theodorou, J. Buchli, S.Schaal AISTATS 2010

E. Theodorou, J. Buchli, S.Schaal JMLR 2010

Path Integral Optimal Control



Cost $J(\mathbf{x}_j(t:N))$

Probability $P(\mathbf{x}_j(t:N)) = \frac{e^{-\frac{1}{\lambda}P(\mathbf{x}_j(t:N))}}{\int e^{-\frac{1}{\lambda}P(\mathbf{x}_j(t:N))}d\mathbf{x}}$

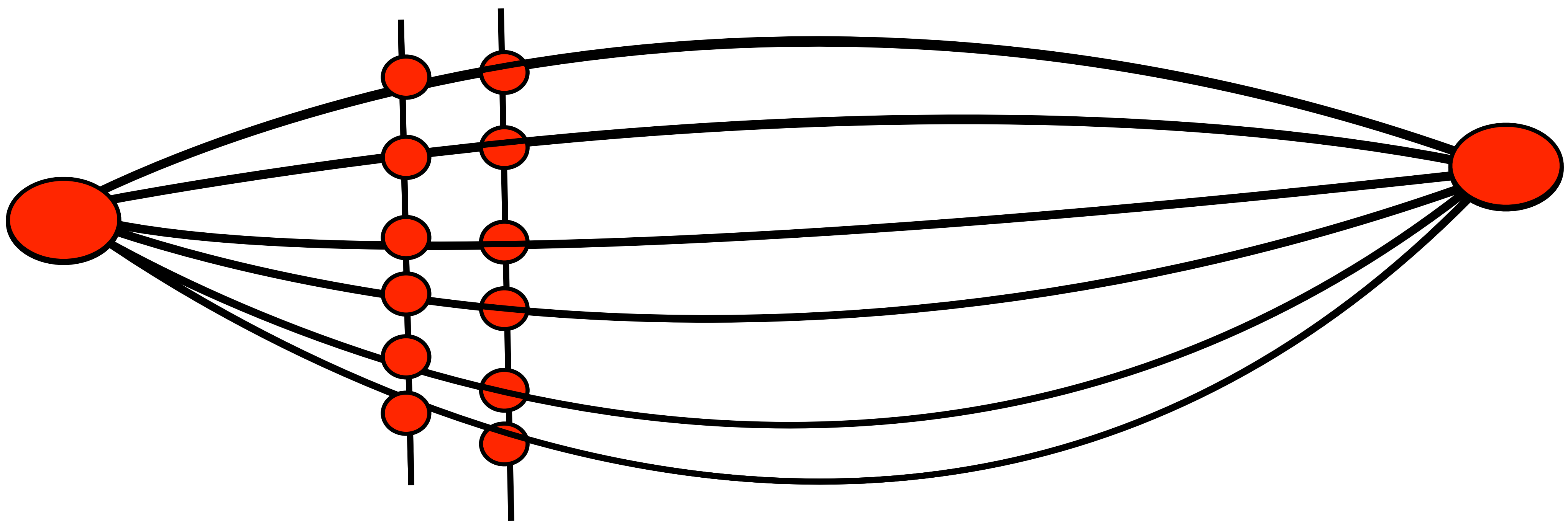
Local controls $\delta\mathbf{u}(t) = \sum_j P(\mathbf{x}_j(t:N))\ \delta\mathbf{u}_j(t)$

E. Theodorou, J. Buchli, S.Schaal ICRA 2010

E. Theodorou, J. Buchli, S.Schaal AISTATS 2010

E. Theodorou, J. Buchli, S.Schaal JMLR 2010

Path Integral Optimal Control



Cost $J(\mathbf{x}_j(t:N))$

Probability $P(\mathbf{x}_j(t:N)) = \dfrac{e^{-\frac{1}{\lambda}P(\mathbf{x}_j(t:N))}}{\int e^{-\frac{1}{\lambda}P(\mathbf{x}_j(t:N))}d\mathbf{x}}$

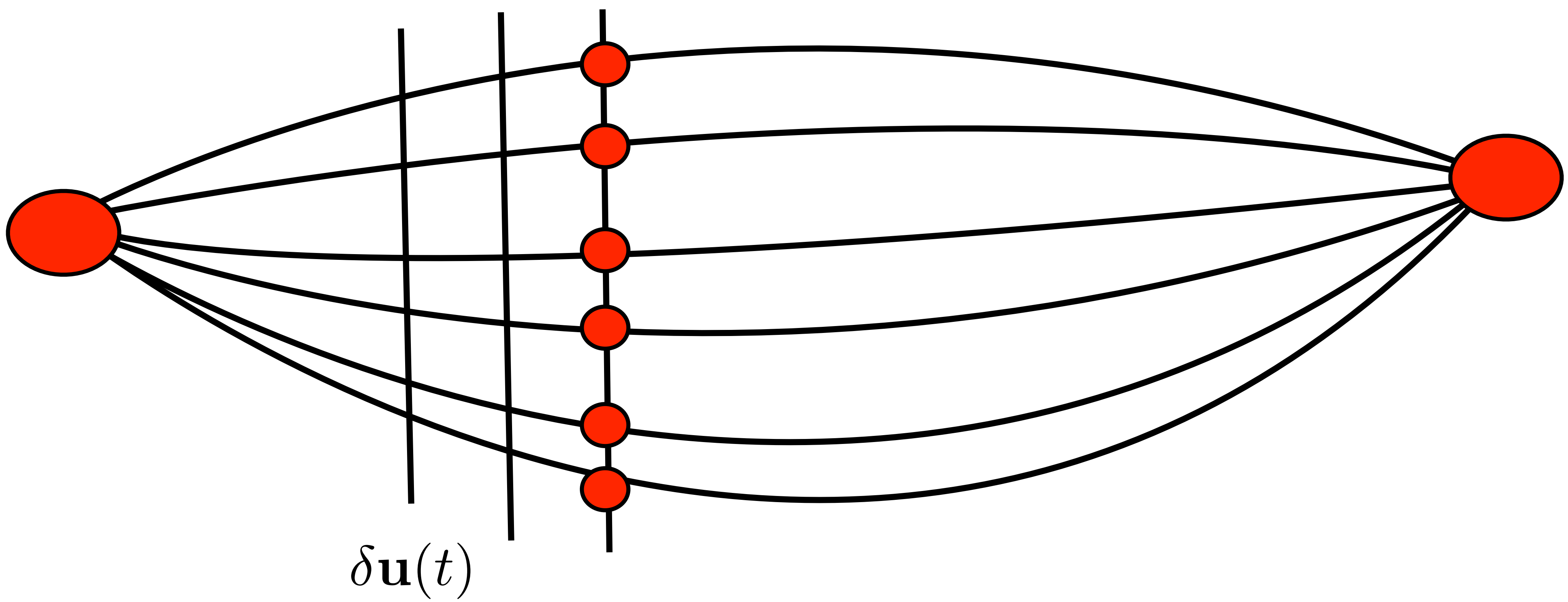
Local controls $\delta \mathbf{u}(t) = \sum_j P(\mathbf{x}_j(t:N))\ \delta \mathbf{u}_j(t)$

E. Theodorou, J. Buchli, S.Schaal ICRA 2010

E. Theodorou, J. Buchli, S.Schaal AISTATS 2010

E. Theodorou, J. Buchli, S.Schaal JMLR 2010

Path Integral Optimal Control



Cost $J(\mathbf{x}_j(t:N))$

Probability $P(\mathbf{x}_j(t:N)) = \dfrac{e^{-\frac{1}{\lambda}P(\mathbf{x}_j(t:N))}}{\int e^{-\frac{1}{\lambda}P(\mathbf{x}_j(t:N))}d\mathbf{x}}$

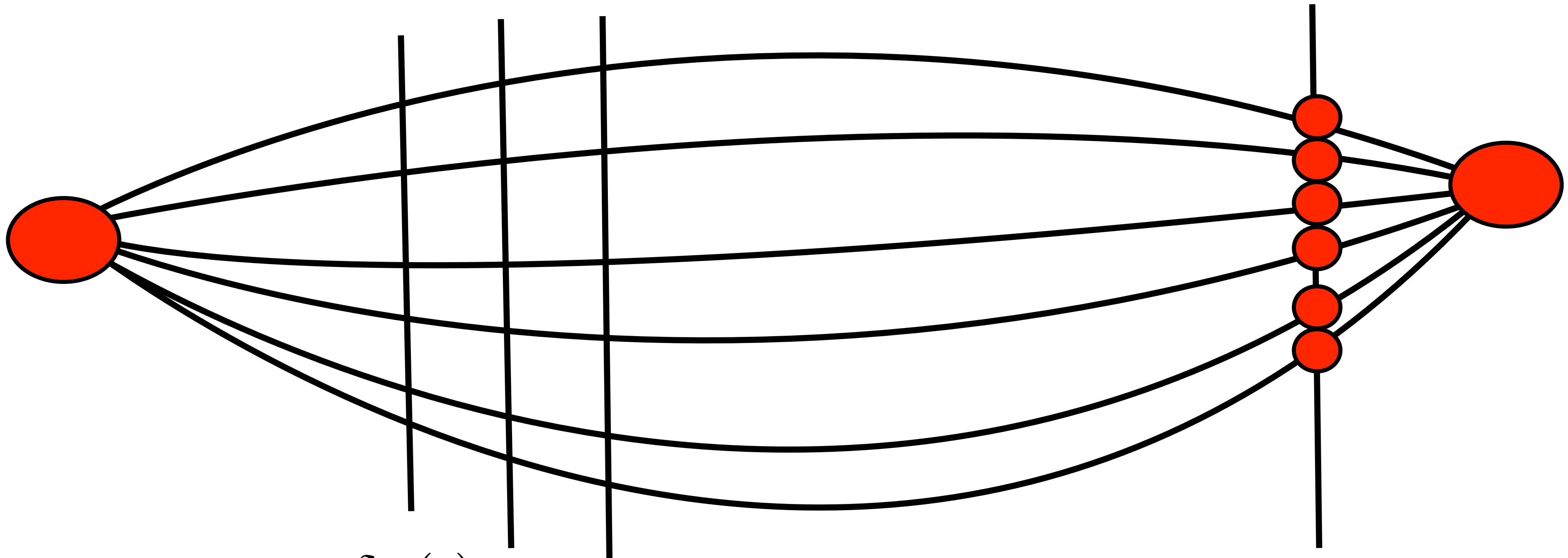
Local controls $\delta\mathbf{u}(t) = \sum_j P(\mathbf{x}_j(t:N))\ \delta\mathbf{u}_j(t)$

E. Theodorou, J. Buchli, S.Schaal ICRA 2010

E. Theodorou, J. Buchli, S.Schaal AISTATS 2010

E. Theodorou, J. Buchli, S.Schaal JMLR 2010

Path Integral Optimal Control



Cost $J(\mathbf{x}_j(t:N))$

Probability $$P(\mathbf{x}_j(t:N)) = \frac{e^{-\frac{1}{\lambda}P(\mathbf{x}_j(t:N))}}{\int e^{-\frac{1}{\lambda}P(\mathbf{x}_j(t:N))}d\mathbf{x}}$$

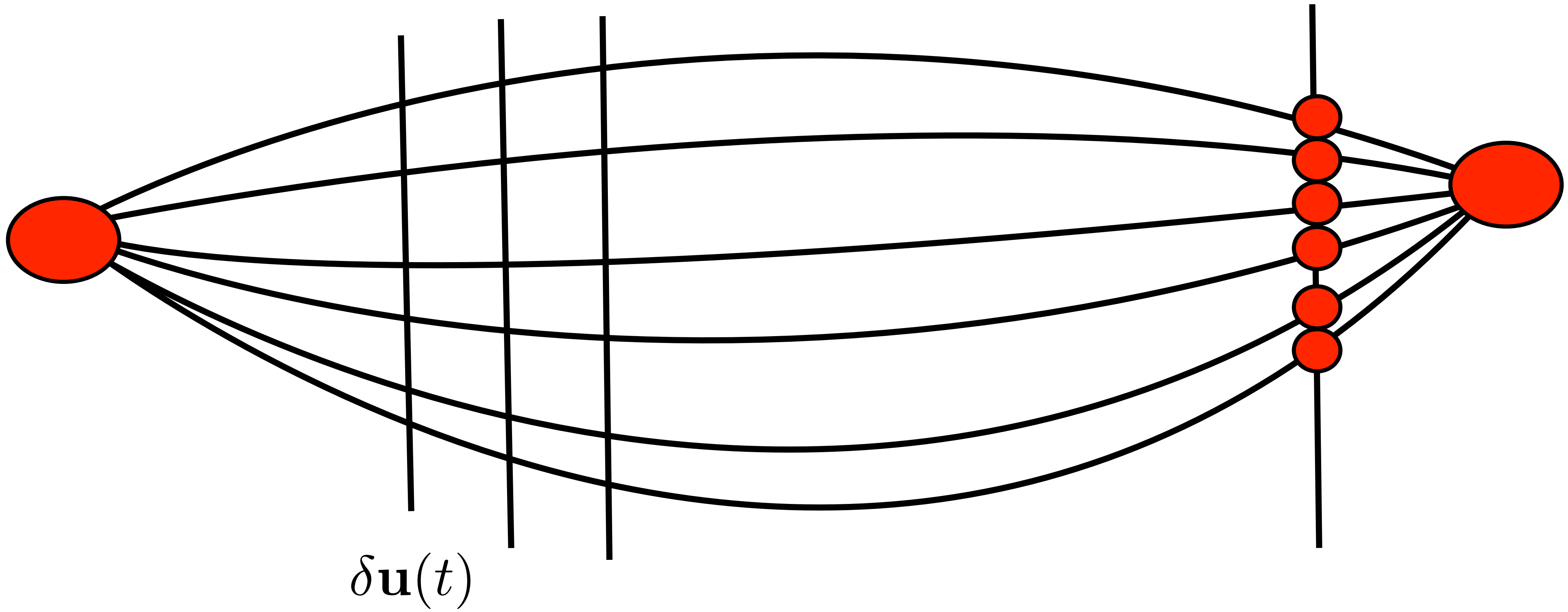
Local controls $$\delta\mathbf{u}(t) = \sum_j P(\mathbf{x}_j(t:N))\ \delta\mathbf{u}_j(t)$$

E. Theodorou, J. Buchli, S.Schaal ICRA 2010

E. Theodorou, J. Buchli, S.Schaal AISTATS 2010

E. Theodorou, J. Buchli, S.Schaal JMLR 2010

Path Integral Optimal Control



Cost $\quad J(\mathbf{x}_j(t:N))$

Probability $\quad P(\mathbf{x}_j(t:N)) = \dfrac{e^{-\frac{1}{\lambda}P(\mathbf{x}_j(t:N))}}{\int e^{-\frac{1}{\lambda}P(\mathbf{x}_j(t:N))}d\mathbf{x}}$

Local controls $\quad \delta\mathbf{u}(t) = \sum_j P(\mathbf{x}_j(t:N))\ \delta\mathbf{u}_j(t)$

$$\mathbf{u}_{new} = \mathbf{u}_{old} + \delta\mathbf{u}$$

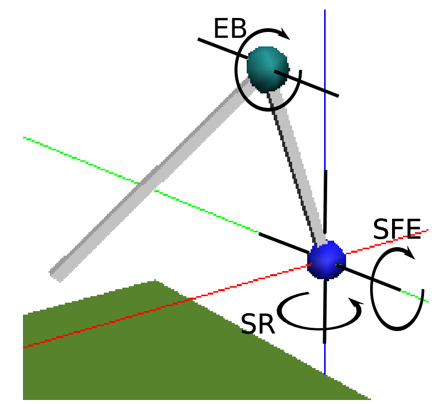
E. Theodorou, J. Buchli, S.Schaal ICRA 2010

E. Theodorou, J. Buchli, S.Schaal AISTATS 2010

E. Theodorou, J. Buchli, S.Schaal JMLR 2010

Path Integral Optimal Control

$$J(\mathbf{x},\mathbf{u},t) = \phi(\mathbf{x}(t_N)) + \int \left(q(\mathbf{x}) + \left|\left|\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right|\right|_p + \left|\left|\frac{\partial \mathbf{u}}{\partial t}\right|\right|_p \right) dt$$



Variable Stiffness Control (Gain Scheduling)

$$J(\mathbf{x},\mathbf{u},t) = \int \left(w_{accel}||\ddot{\mathbf{x}}(t)|| + w_{target}C(t) + w_{gain}\sum_{i}^{\#DOFs} K^{i}_{P,t} \right) dt$$

Acceleration Joint Angles Control Gains

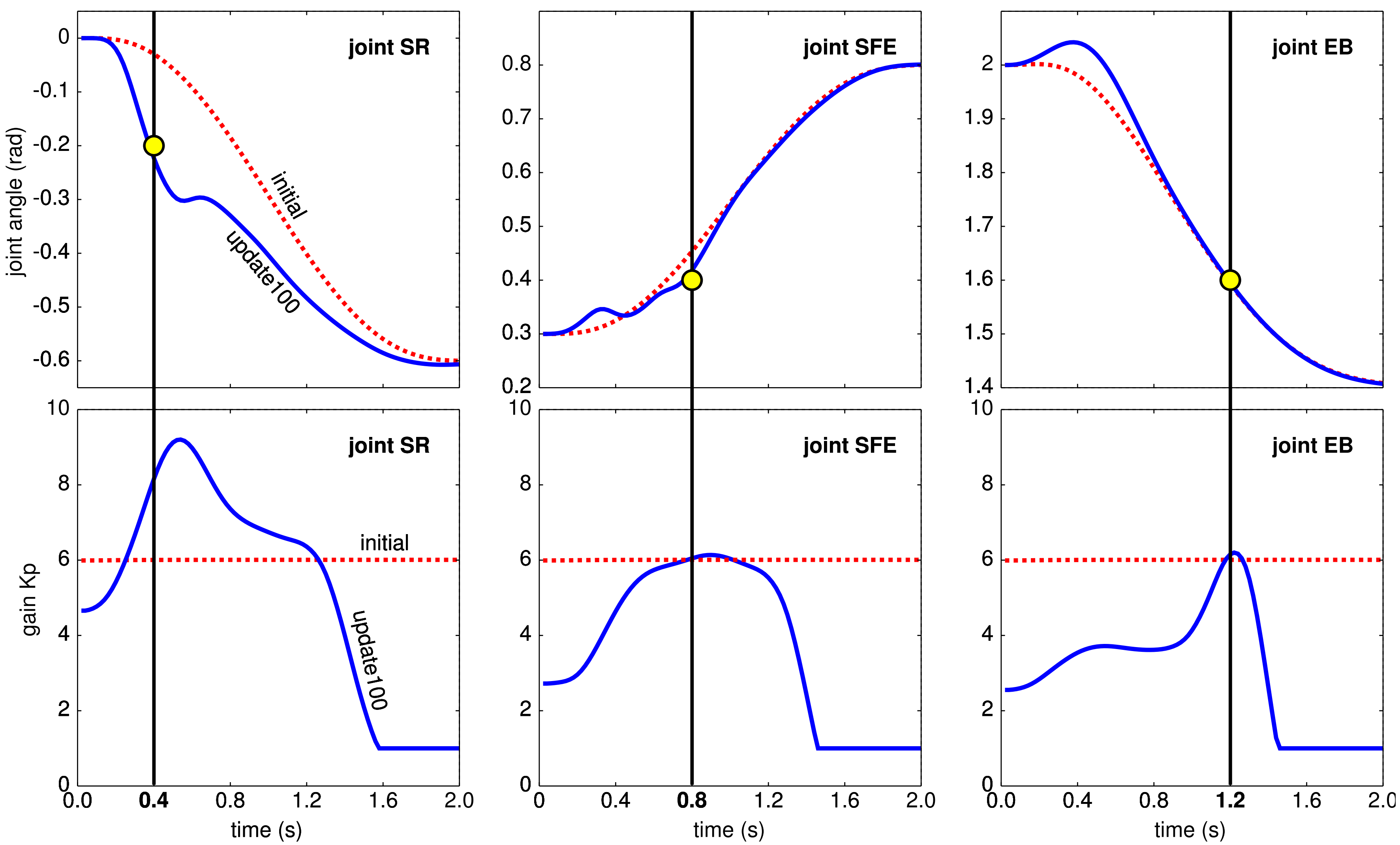
J. Buchli, E.Theodorou, F. Stulp,S. Schaal RSS2010

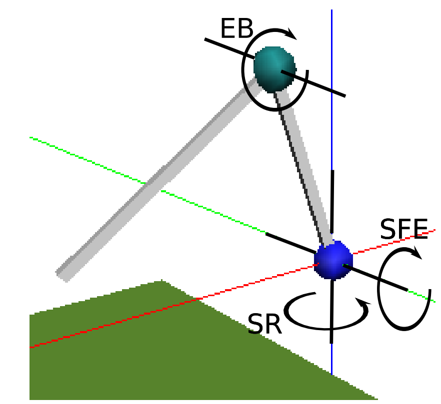
J. Buchli, F. Stulp,E.Theodorou,S. Schaal IJRR 2010

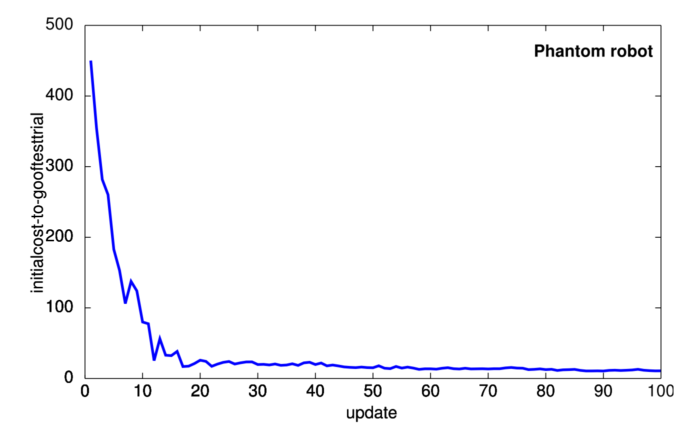
F. Stulp, E.Theodorou, J. Buchli S. Schaal Humanoids 2010

Path Integral Optimal Control

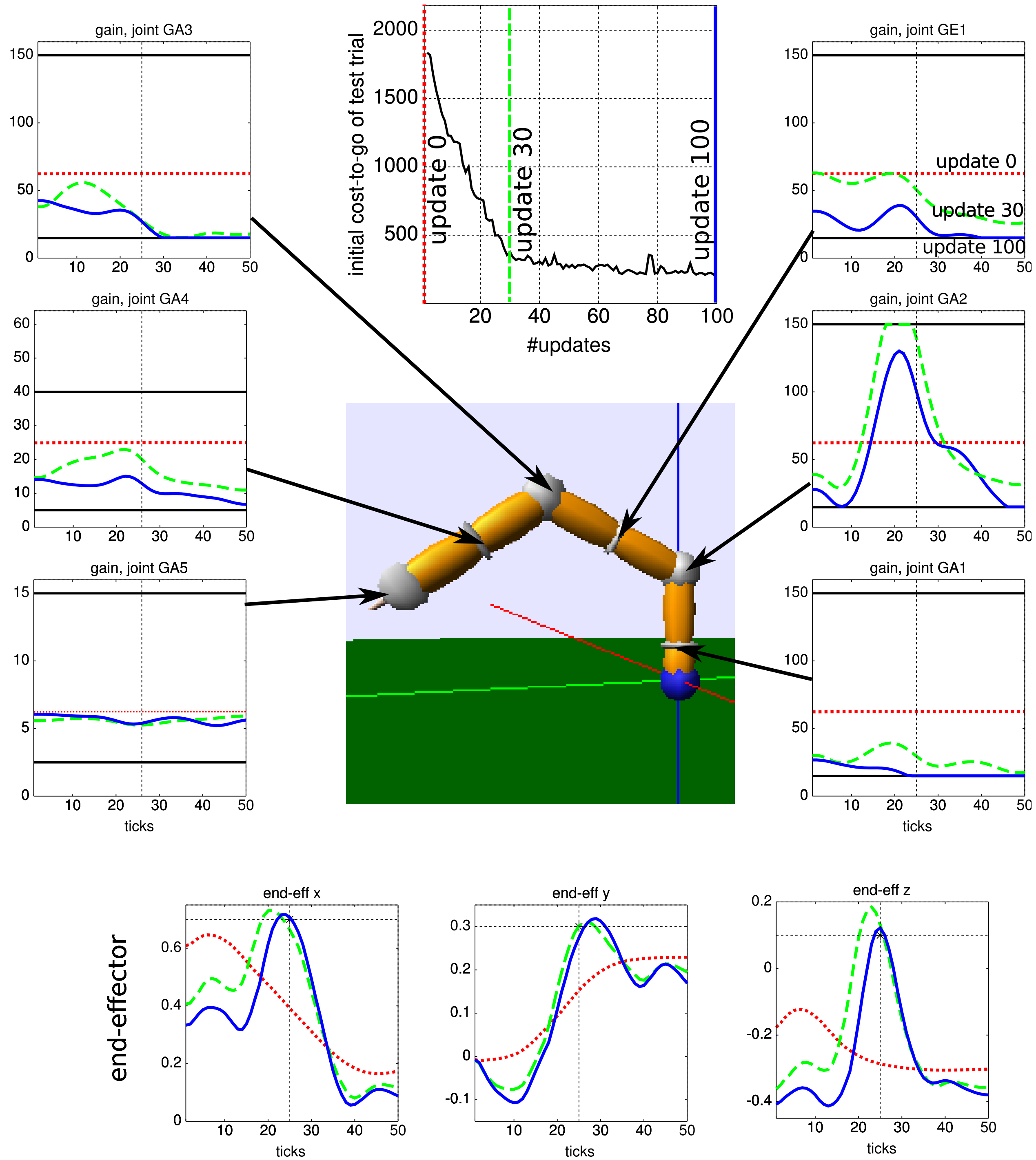
Variable Stiffness Control (Gain Scheduling)



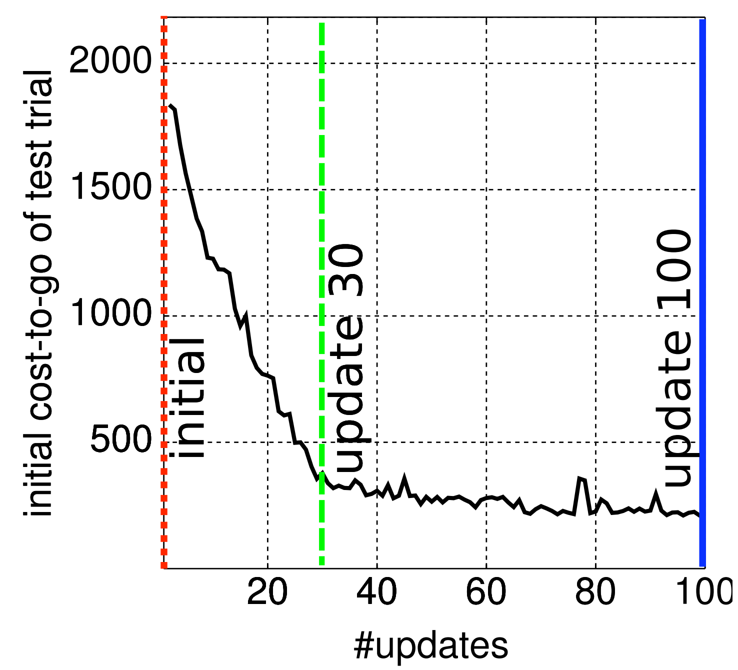




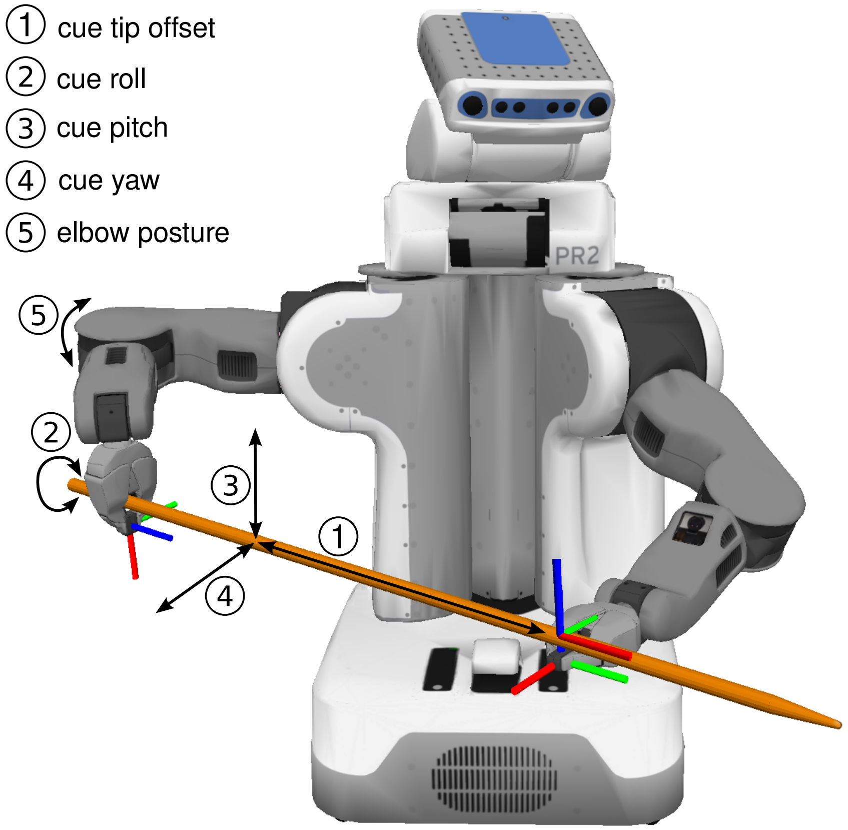
Path Integral Optimal Control



$$C(t) = \delta(t - 0.5)(\mathbf{x} - [\ 0.7\ 0.3\ 0.1]^T)$$



Path Integral Optimal Control

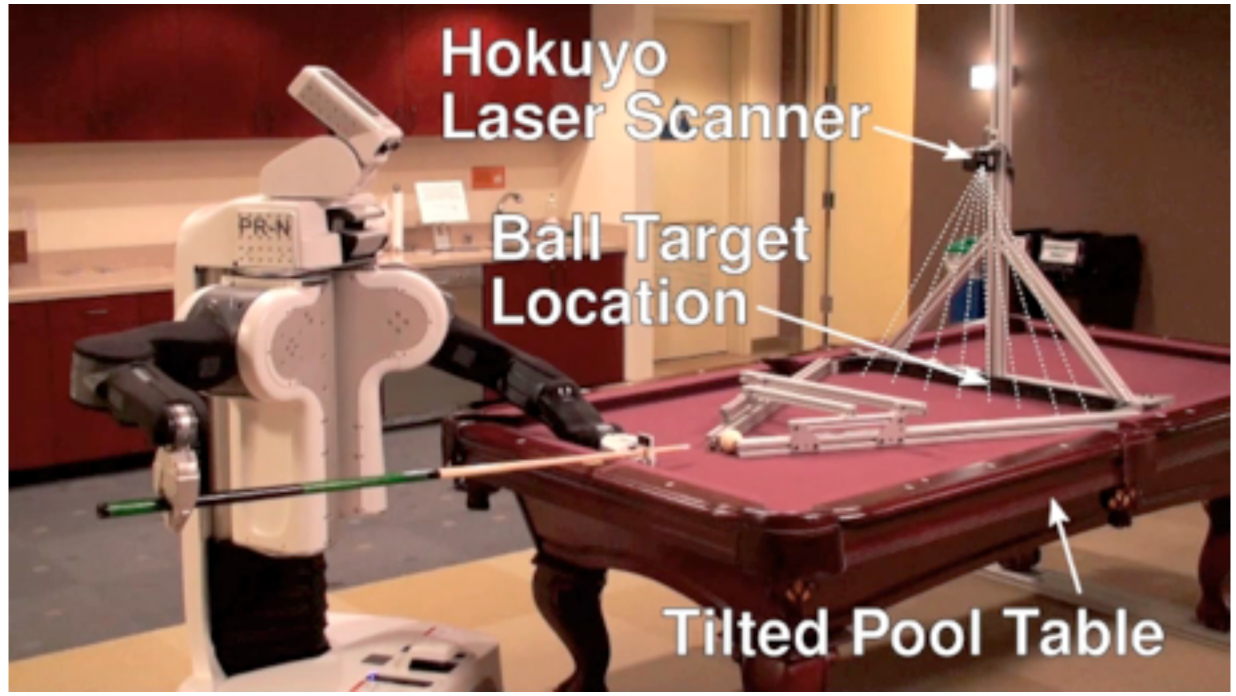


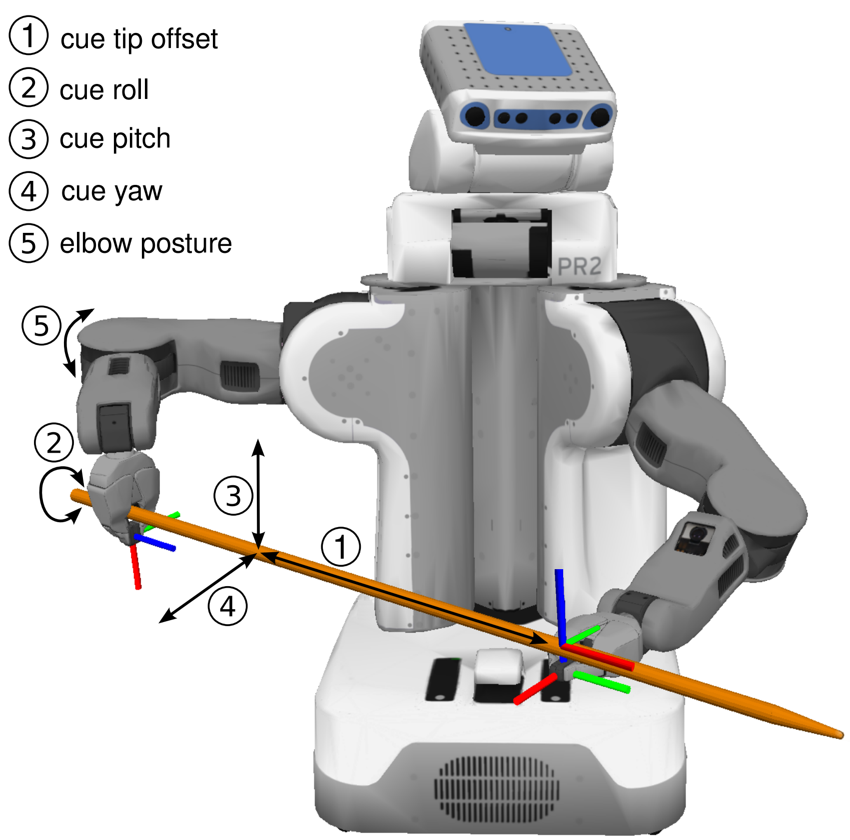
Learning Pool Stroke

$$q(x_t) = w_1 \times Time + w_2 \times Error + w_3 \times Displacement$$

P. Pastor, M. Kalakrishnan, S. Chitta, E.A. Theodorou, S. Schaal Skil, ICRA 2011

Path Integral Optimal Control





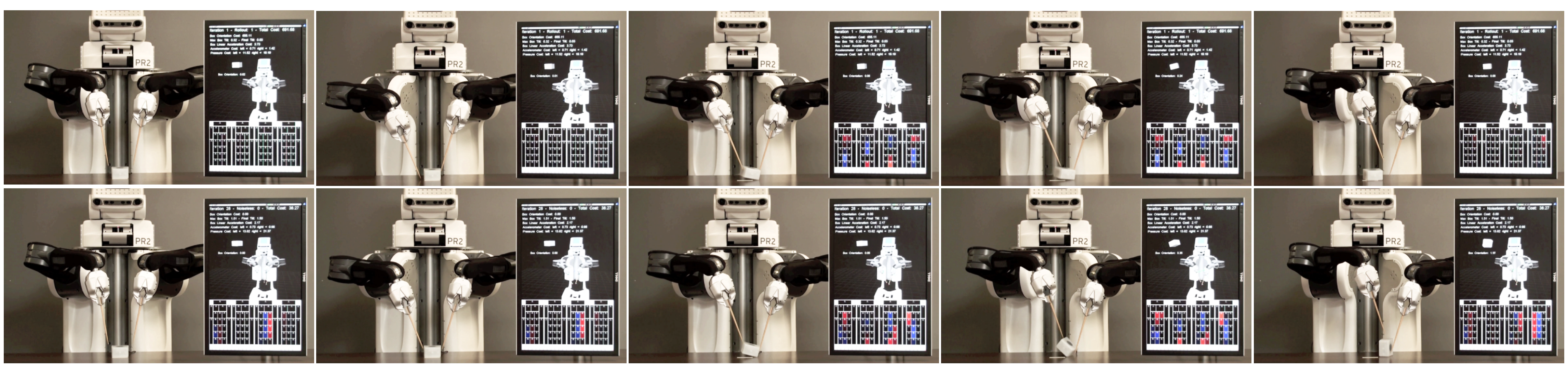
Learning Pool Stroke

$$
\begin{aligned}
q(x_t) &= w_1 \times Time + w_2 \times Error \\
&+ w_3 \times Displacement
\end{aligned}
$$

P. Pastor, M. Kalakrishnan, S. Chitta, E.A. Theodorou, S. Schaal Skil, ICRA 2011

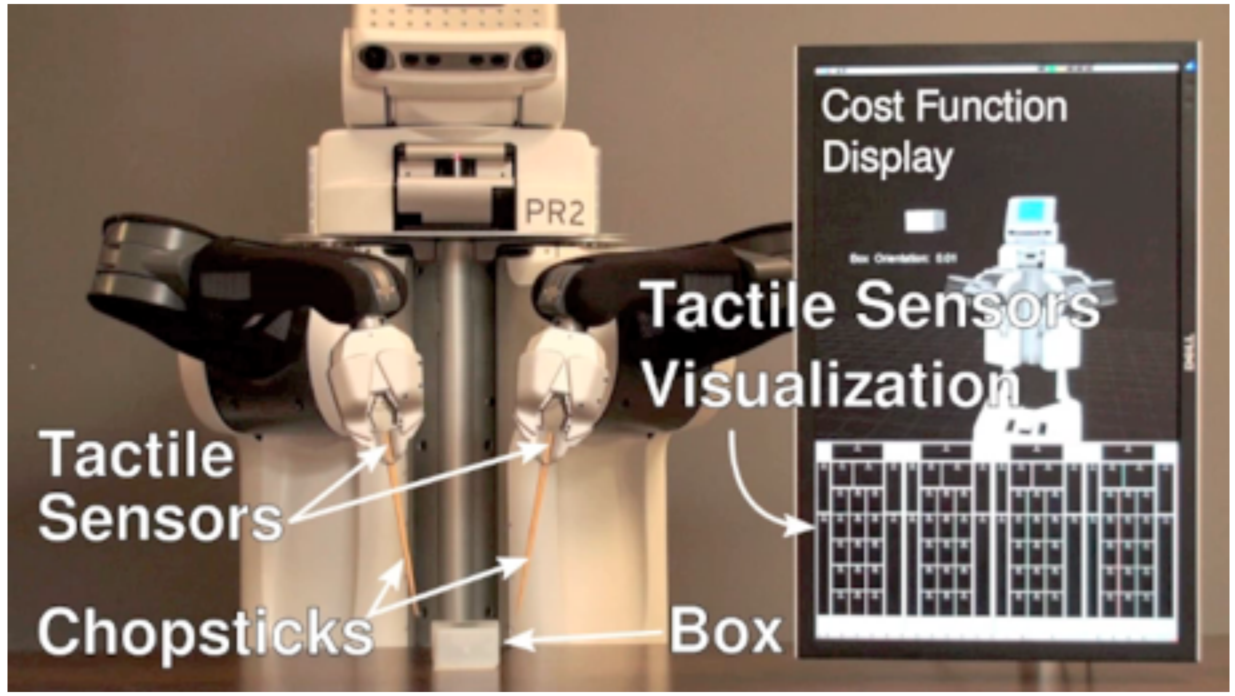
Path Integral Optimal Control

$$C(t) = w_1 \times (\text{Box Acceleration}) + w_2 \times (\text{Force}) + w_3 \times (\text{Gripper acceleration})$$
$$\phi(t) = w_4 \times (\text{Terminal state error})$$

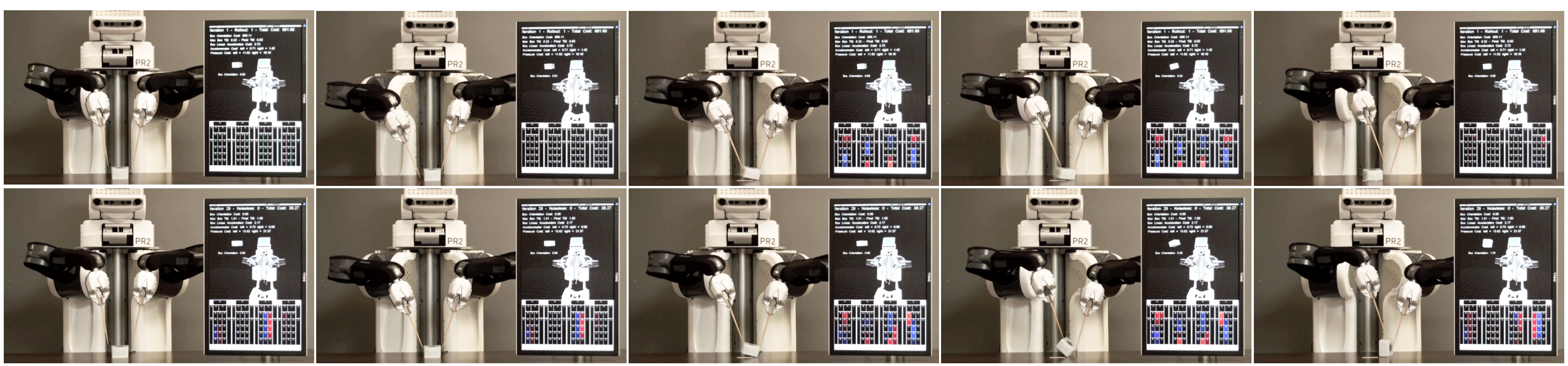


P. Pastor, M. Kalakrishnan, S. Chitta, E.A. Theodorou, S. Schaal Skil, ICRA 2011

Path Integral Optimal Control



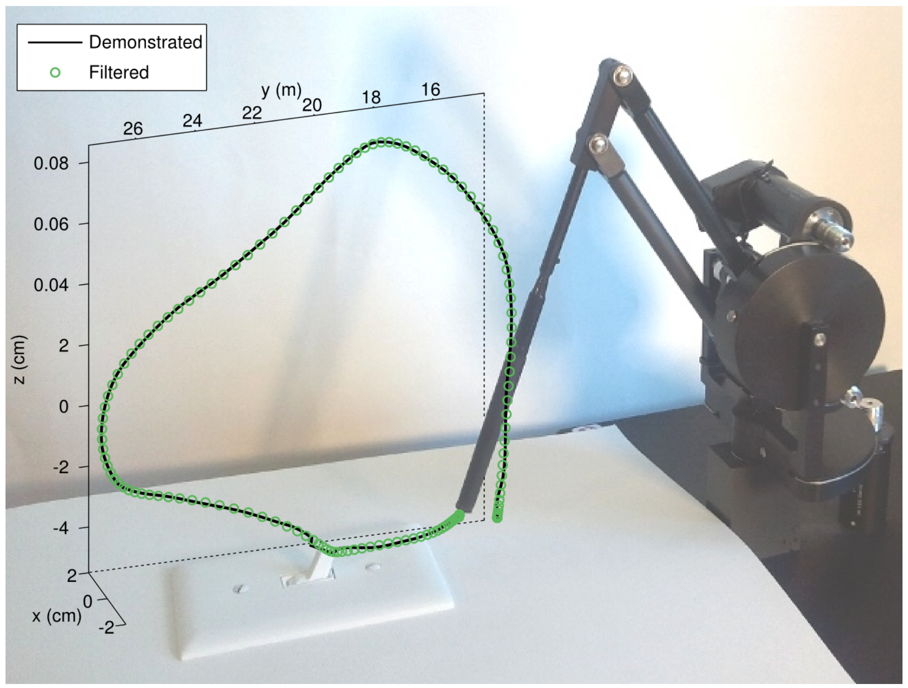
$$C(t) = w_1 \times (\text{Box Acceleration}) + w_2 \times (\text{Force}) + w_3 \times (\text{Gripper acceleration})$$
$$\phi(t) = w_4 \times (\text{Terminal state error})$$

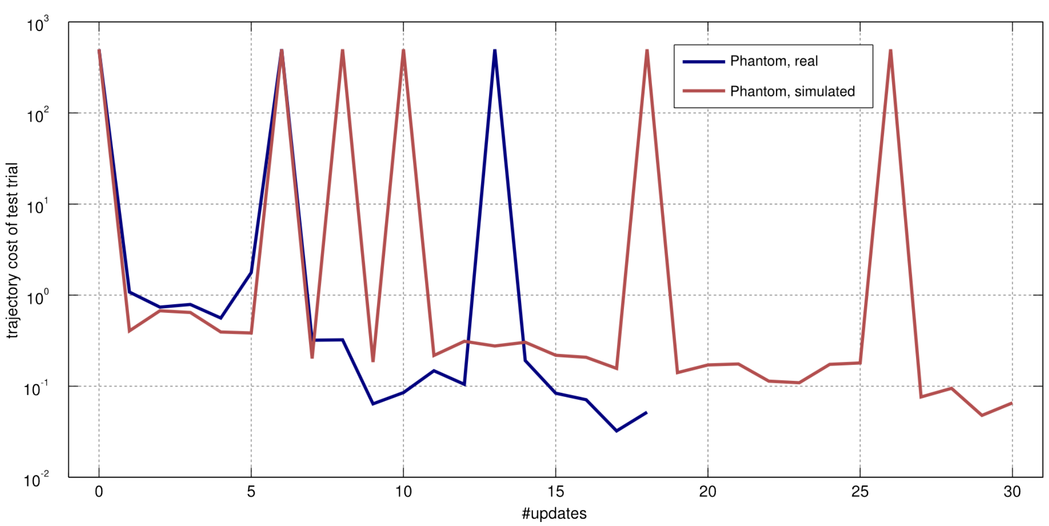


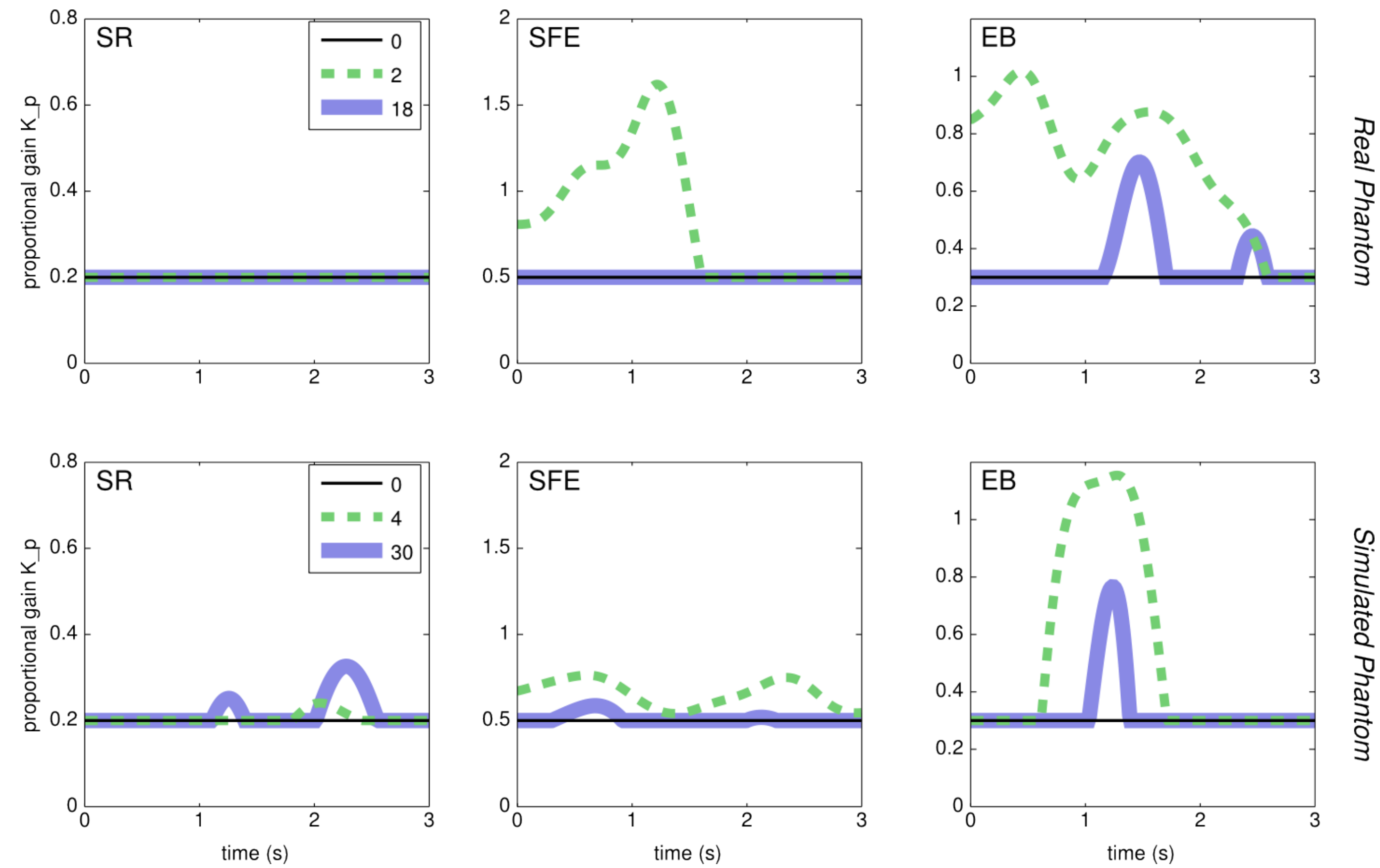
P. Pastor, M. Kalakrishnan, S. Chitta, E.A. Theodorou, S. Schaal Skil, ICRA 2011

Path Integral Optimal Control

Task Goal: Flipping the switch







J. Buchli, F. Stulp,E.Theodorou,S. Schaal IJRR (Accepted)

Path Integral Optimal Control

Jump over the barrier - Little Dog

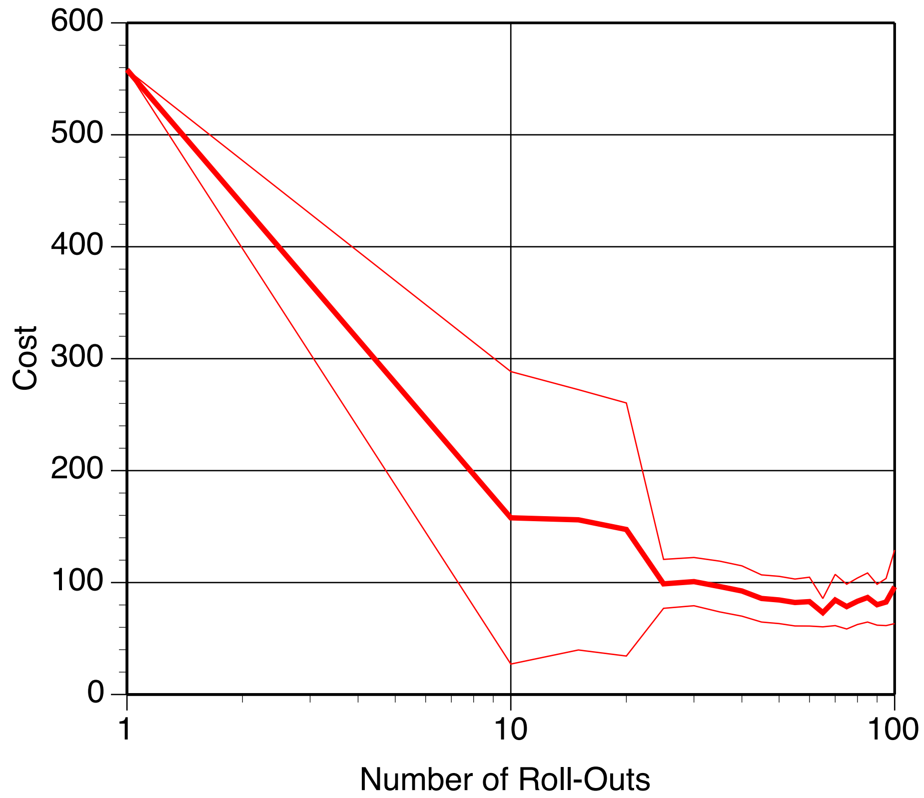
$$r_t = r_{roll} + r_{yaw} + \sum_{i=1}^{d} \left(a_1 f_{i,t}^2 + 0.5 a_2 \; \boldsymbol{\theta}_i^T \boldsymbol{\theta} \right) \; (a_1 = 1.e - 6, a_2 = 1.e - 8)$$

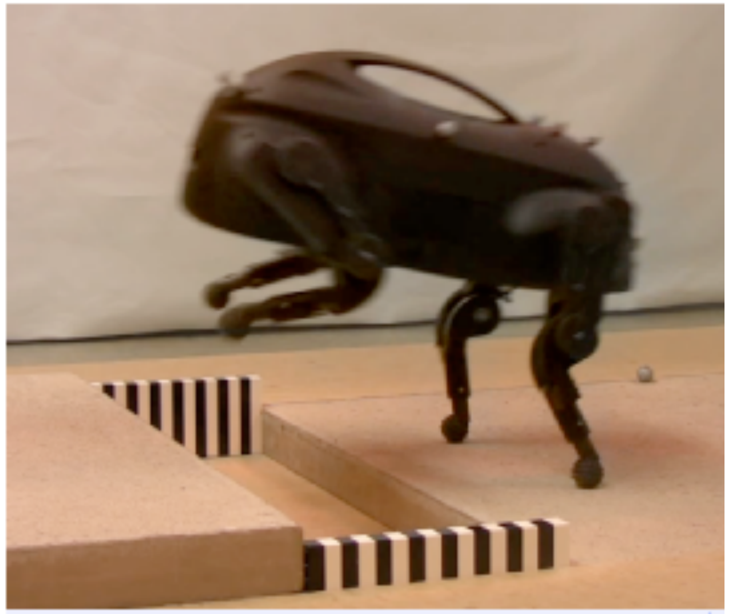
$$r_{roll} = \begin{cases} 100 * (|roll_t| - 0.3)^2, & \text{if } (|roll_t| > 0.3) \\ 0, & otherwise \end{cases}$$

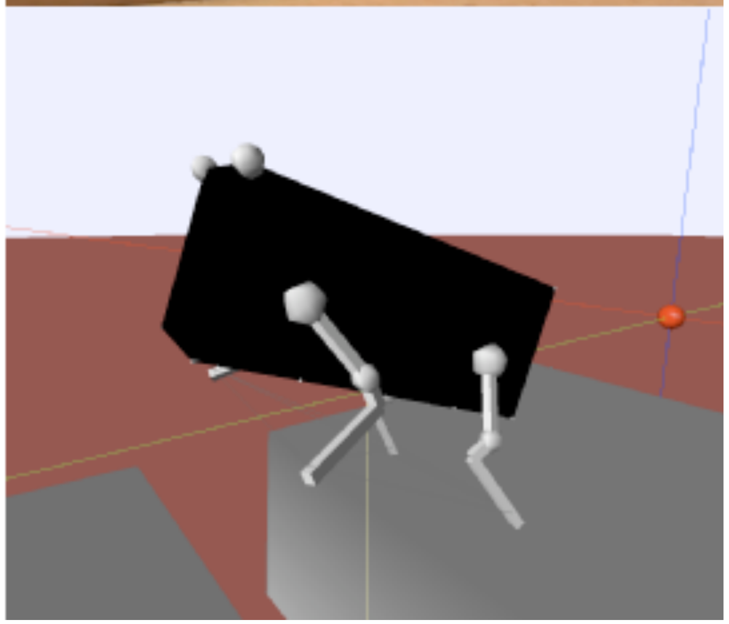
$$r_{yaw} = \begin{cases} 100 * (|yaw_t| - 0.1)^2, & \text{if } (|yaw_t| > 0.1) \\ 0, & otherwise \end{cases}$$

$$\phi_{t_N} = 50000(goal - x_{nose})^2$$

Path Integral Optimal Control







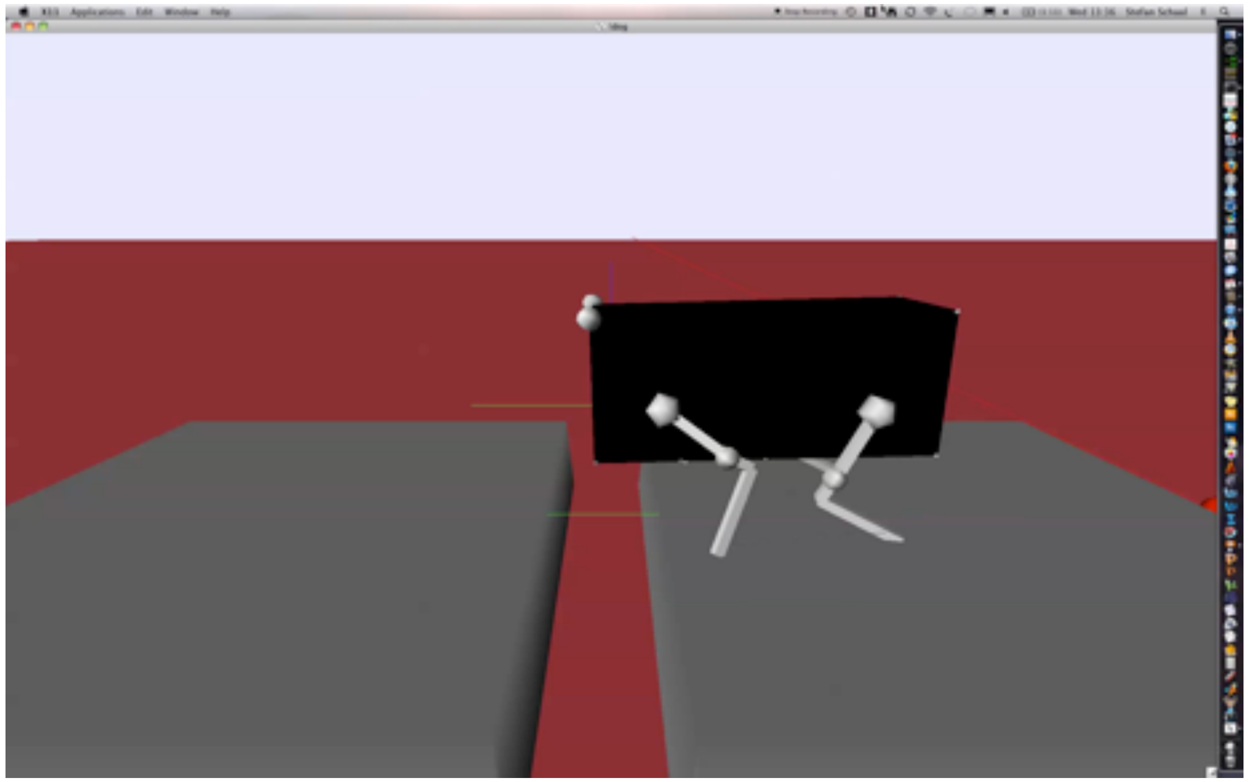
Path Integral Optimal Control

Before Learning

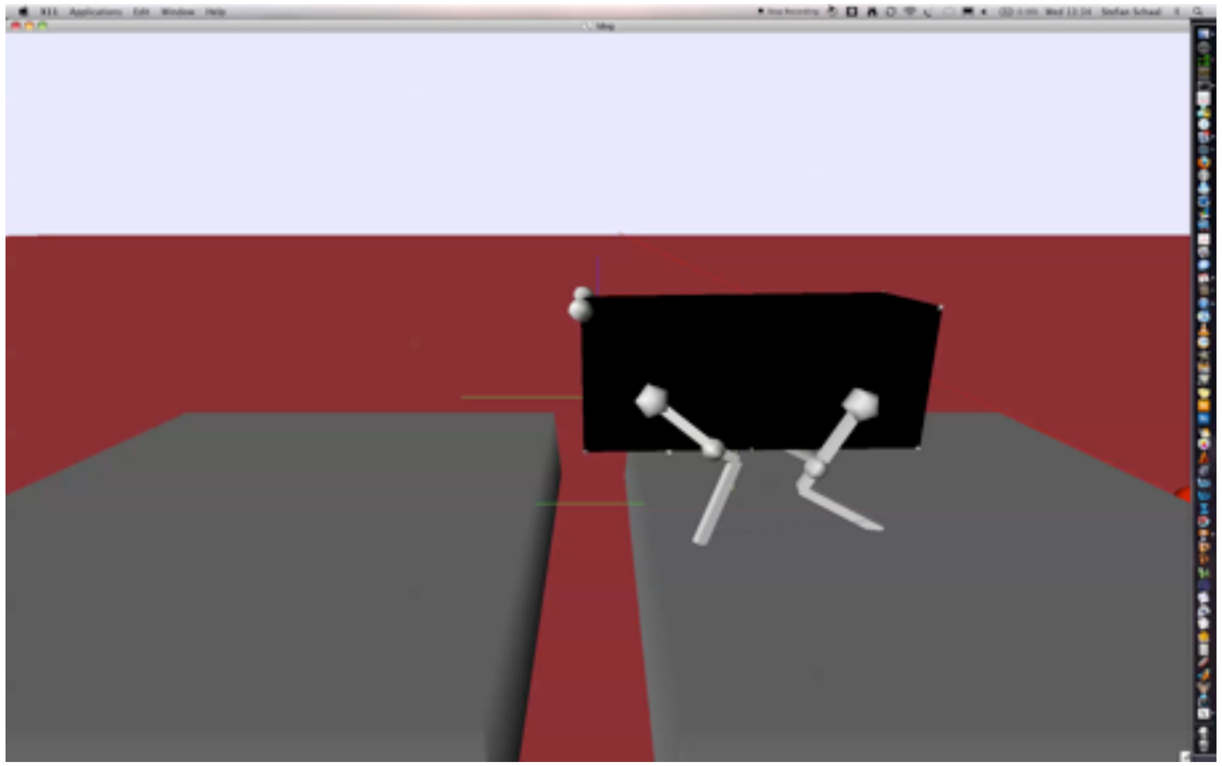
After Learning

Path Integral Optimal Control

Before Learning

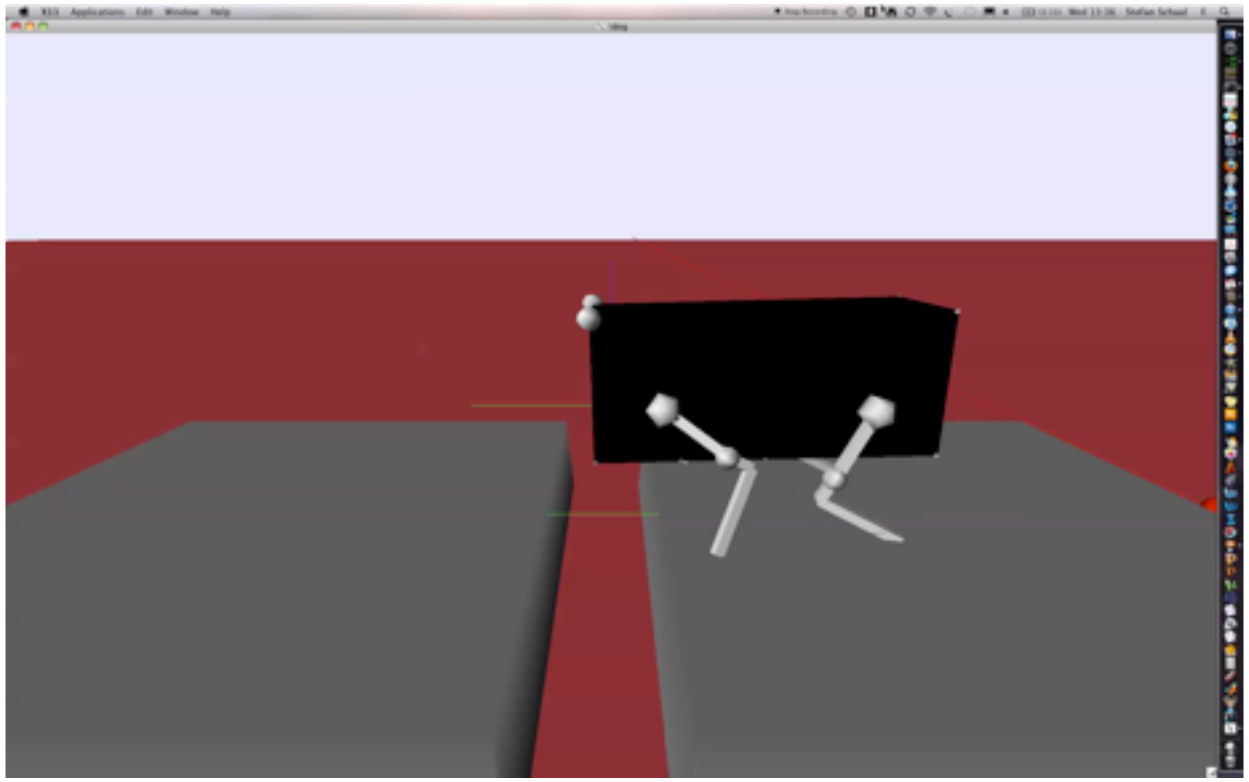


After Learning

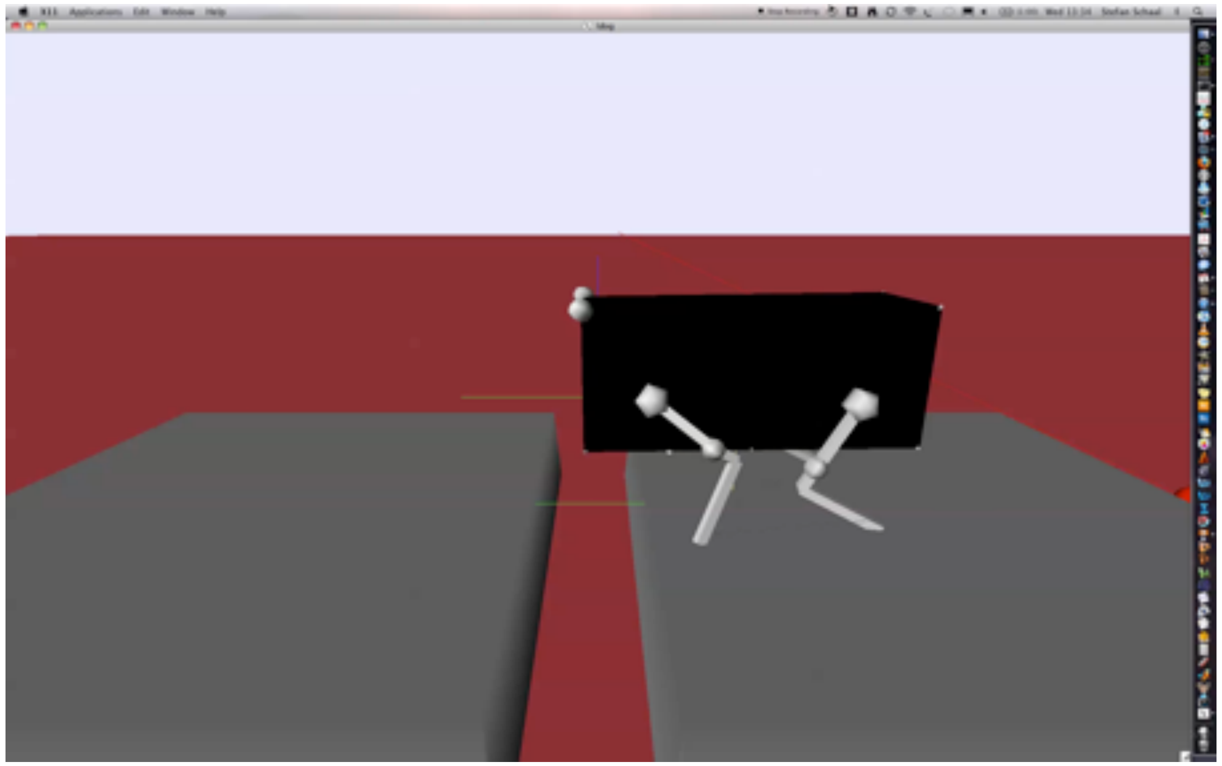


Path Integral Optimal Control

Before Learning



After Learning



Part 2
Beyond Linearization and internal models: Path Integral Optimal Control

Advantages!

Part 2
Beyond Linearization and internal models: Path Integral Optimal Control

Advantages!

1) Model Free & Model Based

Part 2
Beyond Linearization and internal models: Path Integral Optimal Control

Advantages!

1) Model Free & Model Based

2) Forward problem

Part 2
Beyond Linearization and internal models: Path Integral Optimal Control

Advantages!

1) Model Free & Model Based

2) Forward problem

3) Probabilistic - Fast

Part 2
Beyond Linearization and internal models: Path Integral Optimal Control

Advantages!

1) Model Free & Model Based

2) Forward problem

3) Probabilistic - Fast

4) Robust Performance in High Dimensional robotic control problem

Part 2
Beyond Linearization and internal models: Path Integral Optimal Control

Advantages!

1) Model Free & Model Based

2) Forward problem

3) Probabilistic - Fast

4) Robust Performance in High Dimensional robotic control problem

5) It works on physical systems, not only in simulation!

Part 2
Beyond Linearization and internal models: Path Integral Optimal Control

Advantages!

1) Model Free & Model Based

2) Forward problem

3) Probabilistic - Fast

4) Robust Performance in High Dimensional robotic control problem

5) It works on physical systems, not only in simulation!

6) it may actually recapitulate the sensorimotor learning process...

Part 2
Beyond Linearization and internal models: Path Integral Optimal Control

USC VITERBI SCHOOL OF ENGINEERING

Advantages!

1) Model Free & Model Based

2) Forward problem

3) Probabilistic - Fast

4) Robust Performance in High Dimensional robotic control problem

5) It works on physical systems, not only in simulation!

6) it may actually recapitulate the sensorimotor learning process...

Be aware!

Part 2
Beyond Linearization and internal models: Path Integral Optimal Control

Advantages!

1) Model Free & Model Based

2) Forward problem

3) Probabilistic - Fast

4) Robust Performance in High Dimensional robotic control problem

5) It works on physical systems, not only in simulation!

6) it may actually recapitulate the sensorimotor learning process...

Be aware!

1) Initial set of trajectories are required - Learning from demonstration

Part 2
Beyond Linearization and internal models: Path Integral Optimal Control

Advantages!

1) Model Free & Model Based

2) Forward problem

3) Probabilistic - Fast

4) Robust Performance in High Dimensional robotic control problem

5) It works on physical systems, not only in simulation!

6) it may actually recapitulate the sensorimotor learning process...

Be aware!

1) Initial set of trajectories are required - Learning from demonstration

2) An initial stabilizing control policy

General comments about control under uncertainty.

General comments about control under uncertainty.

Sensitivity and robustness to plant uncertainty or model errors.

General comments about control under uncertainty.

Sensitivity and robustness to plant uncertainty or model errors.

Stochastic optimal control has poor robustness and stability margins J Doyle 1972, M. Safonov 1971.

General comments about control under uncertainty.

Sensitivity and robustness to plant uncertainty or model errors.

Stochastic optimal control has poor robustness and stability margins J Doyle 1972, M. Safonov 1971.

This is important for biological motor control if one considers the parametric variability in neuromuscular models.

General comments about control under uncertainty.

Sensitivity and robustness to plant uncertainty or model errors.

Stochastic optimal control has poor robustness and stability margins J Doyle 1972, M. Safonov 1971.

This is important for biological motor control if one considers the parametric variability in neuromuscular models.

Control under uncertainty:

General comments about control under uncertainty.

Sensitivity and robustness to plant uncertainty or model errors.

Stochastic optimal control has poor robustness and stability margins J Doyle 1972, M. Safonov 1971.

This is important for biological motor control if one considers the parametric variability in neuromuscular models.

Control under uncertainty:

Model Predictive	Re-optimization on the fly	Heavy Computations	

General comments about control under uncertainty.

Sensitivity and robustness to plant uncertainty or model errors.

Stochastic optimal control has poor robustness and stability margins J Doyle 1972, M. Safonov 1971.

This is important for biological motor control if one considers the parametric variability in neuromuscular models.

Control under uncertainty:

Model Predictive	Re-optimization on the fly	Heavy Computations	
Robust Control	Bounded Disturbances (small gain theorem)	Conservative	

General comments about control under uncertainty.

Sensitivity and robustness to plant uncertainty or model errors.

Stochastic optimal control has poor robustness and stability margins J Doyle 1972, M. Safonov 1971.

This is important for biological motor control if one considers the parametric variability in neuromuscular models.

Control under uncertainty:

Model Predictive	Re-optimization on the fly	Heavy Computations	
Robust Control	Bounded Disturbances (small gain theorem)	Conservative	
Adaptive Control	Unknown parameters Adaptive law, PE.	Sensitivity stochasticity Observation noise	

General comments about control under uncertainty.

Sensitivity and robustness to plant uncertainty or model errors.

Stochastic optimal control has poor robustness and stability margins J Doyle 1972, M. Safonov 1971.

This is important for biological motor control if one considers the parametric variability in neuromuscular models.

Control under uncertainty:

Model Predictive	Re-optimization on the fly	Heavy Computations	
Robust Control	Bounded Disturbances (small gain theorem)	Conservative	
Adaptive Control	Unknown parameters Adaptive law, PE.	Sensitivity stochasticity Observation noise	
Model Free OC or RL	Learning from experience	Initial Policy or demonstration	

General comments about control under uncertainty.

Sensitivity and robustness to plant uncertainty or model errors.

Stochastic optimal control has poor robustness and stability margins J Doyle 1972, M. Safonov 1971.

This is important for biological motor control if one considers the parametric variability in neuromuscular models.

Control under uncertainty:

Model Predictive	Re-optimization on the fly	Heavy Computations	Internal Models
Robust Control	Bounded Disturbances (small gain theorem)	Conservative	
Adaptive Control	Unknown parameters Adaptive law, PE.	Sensitivity stochasticity Observation noise	
Model Free OC or RL	Learning from experience	Initial Policy or demonstration	

Acknowledgments

PIs

Stefan Schaal

FJ Valero Cuevas

Emannuel Todorov

J. Buchli

Freek Stulp

P. Pastor

M. Kalakrishnan

M. Manish

J. Kutch

Y. Tassa