EP for Efficient Stochastic Control with Obstacles

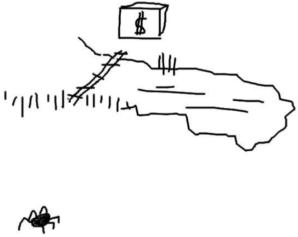
Thomas Mensink[1] Jakob Verbeek[1] Bert Kappen[2]

LEAR - INRIA Rhône-Alpes, Grenoble, France
SNN - Radboud University, Nijmegen, The Netherlands

European Conference on Artificial Intelligence - Lisbon Portugal
August 19th, 2010

Introduction

- Optimal control is how to act now to optimize future rewards
- Real life systems have constraints on the allowed state
  - angle of joint of a robot arm
  - part of the road used by an autonomous car
- The presence of (Wiener) noise, makes stochastic control quantitatively different from deterministic control



Overview

Path Integral (1)

- Class of continuous non-linear control problems, written as Path Integral over a forward diffusion process
- Stochastic dynamical system

$$dx = (f(t,x) + B\,u)\,dt + d\xi \tag{1}$$

$$C(t,x,u(t \to t_f)) = \left\langle \underbrace{\phi(x_{t_f})}_{\text{end cost}} + \int_t^{t_f} dt \; \underbrace{\frac{1}{2} u_t^\top R u_t}_{\text{control cost}} + \underbrace{V(x_t,t)}_{\text{state cost}} \right\rangle \tag{2}$$

- The forward path integral for $t < t_f$, is

$$\psi(x_t,t) = \int dy \rho(y,t_f|x_t,t)\psi(y,t_f), \qquad \psi(y,t_f) = \exp(-\phi(y)/\lambda). \tag{3}$$

- Diffusion process ρ is a Fokker-Planck equation with drift $f(x_t,t)\,dt$ and diffusion $d\xi$, and an term due to $V(x_t,t)$.

Path Integral (1)

- Class of continuous non-linear control problems, written as Path Integral over a forward diffusion process
- Stochastic dynamical system

$$dx = (f(t,x) + B\,u)\,dt + d\xi \tag{1}$$

$$C(t,x,u(t \to t_f)) = \left\langle \underbrace{\phi(x_{t_f})}_{\text{end cost}} + \int_t^{t_f} dt \;\; \underbrace{\frac{1}{2} u_t^\top R u_t}_{\text{control cost}} + \underbrace{V(x_t,t)}_{\text{state cost}} \right\rangle \tag{2}$$

- The forward path integral for $t < t_f$, is

$$\psi(x_t,t) = \int dy \rho(y,t_f|x_t,t)\psi(y,t_f), \qquad \psi(y,t_f) = \exp(-\phi(y)/\lambda). \tag{3}$$

- Diffusion process ρ is a Fokker-Planck equation with drift $f(x_t,t)\,dt$ and diffusion $d\xi$, and an term due to $V(x_t,t)$.

Path Integral (1)

- Class of continuous non-linear control problems, written as Path Integral over a forward diffusion process
- Stochastic dynamical system

$$dx = (f(t,x) + B\,u)\,dt + d\xi \tag{1}$$

$$C(t,x,u(t \to t_f)) = \left\langle \underbrace{\phi(x_{t_f})}_{\text{end cost}} + \int_t^{t_f} dt \;\; \underbrace{\frac{1}{2}u_t^\top R u_t}_{\text{control cost}} + \underbrace{V(x_t,t)}_{\text{state cost}} \right\rangle \tag{2}$$

- The forward path integral for $t < t_f$, is

$$\psi(x_t,t) = \int dy \rho(y,t_f|x_t,t)\psi(y,t_f), \qquad \psi(y,t_f) = \exp(-\phi(y)/\lambda). \tag{3}$$

- Diffusion process ρ is a Fokker-Planck equation with drift $f(x_t,t)\,dt$ and diffusion $d\xi$, and an term due to $V(x_t,t)$.

Path Integral (2)

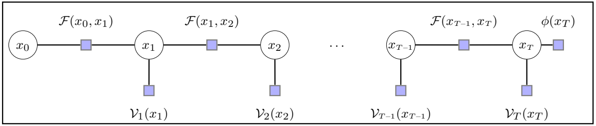
- Path integral can be interpreted as a free energy, therefore approximated with
  - MC sampling,
  - Variational Methods, and
  - Expectation Propagation (EP).

- Promising for multi-agent control and coordination problems and several robot tasks.

- In this presentation: hard walls (interval constraints) as obstacles.

Approximations for Optimal Control

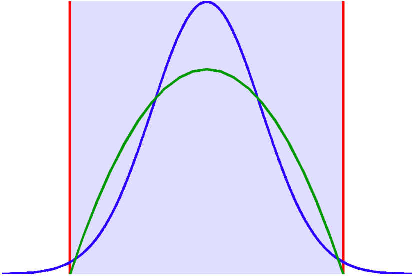
- Forward Diffusion Process



$$P(\mathbf{x}_{1:T}|x_0) = \prod_{t=1}^{T} \mathcal{V}_t(x_t)\, \mathcal{F}(x_t, x_{t-1})\, \phi(x_T) \quad (4)$$

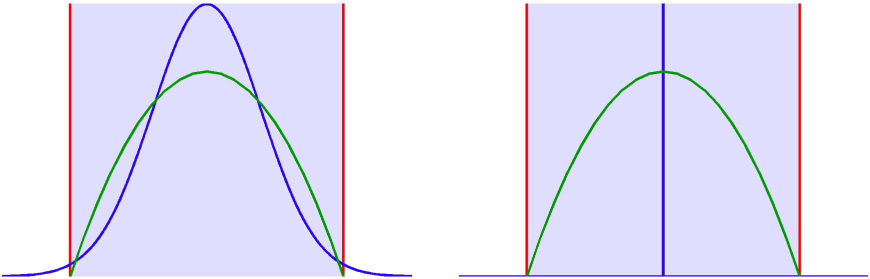
- Approximated with:
  - Variational Method (minimise $\mathrm{KL}(q\|p)$)
  - Expectation Propagation (minimise $\mathrm{KL}(p\|q)$)

Variational Methods vs EP



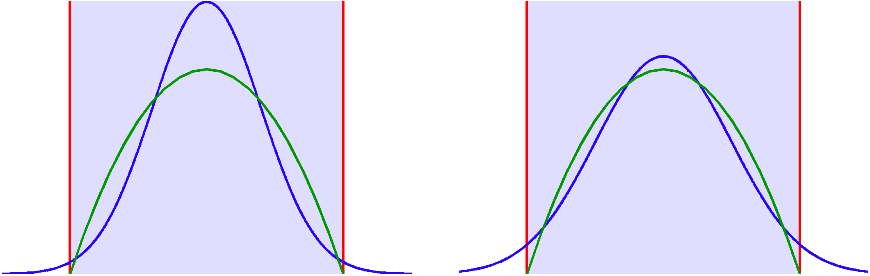
- Variational Methods
  - minimise $\mathrm{KL}(q\|p)$
  - q can not have any probability mass where $p = 0$.
- Expectation Propagation
  - minimise $\mathrm{KL}(p\|q)$
  - q matches the moments of p.

Variational Methods vs EP



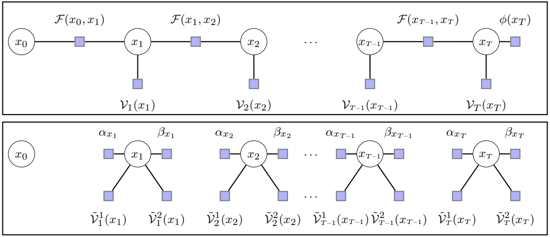
- Variational Methods
  - minimise $\mathrm{KL}(q\|p)$
  - q can not have any probability mass where $p = 0$.
- Expectation Propagation
  - minimise $\mathrm{KL}(p\|q)$
  - q matches the moments of p.

Variational Methods vs EP



- Variational Methods
  - minimise $\mathrm{KL}(q\|p)$
  - q can not have any probability mass where $p = 0$.
- Expectation Propagation
  - minimise $\mathrm{KL}(p\|q)$
  - q matches the moments of p.

Approximation with EP

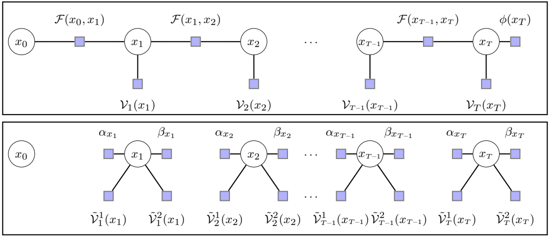


$$Q(\mathbf{x}_{1:T}|x_0) = \prod_{t=1}^{T} \alpha(x_t)\beta(x_t) \prod_{d=1}^{D} \tilde{\mathcal{V}}_t^d(x_t)$$

- Iteratively minimise a factor, taking into account the context.
- Moment matching:

$$q_i' \propto \mathrm{Proj}\left[q^{\setminus i}\, p_i\right] = \min \mathrm{KL}(P \| q^{\setminus i}\, p_i)$$

Approximation with EP

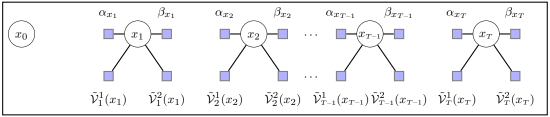


$$Q(\mathbf{x}_{1:T}|x_0) = \prod_{t=1}^{T} \alpha(x_t)\beta(x_t) \prod_{d=1}^{D} \tilde{\mathcal{V}}_t^d(x_t)$$

- Iteratively minimise a factor, taking into account the context.
- Moment matching:

$$q_i' \propto \mathrm{Proj}\left[q^{\setminus i}\, p_i\right] = \min \mathrm{KL}(P \| q^{\setminus i}\, p_i)$$

EP for Graphs - Updating α and β

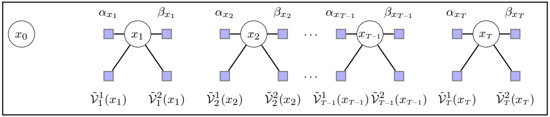


$$q'_{\alpha_t}(x_t) = \text{Proj}\left[q^{\setminus \alpha_t}(x_t) \int \mathcal{F}(x_{t\text{-}1}, x_t)\alpha(x_{t\text{-}1}) \prod_{d=1}^{D} \tilde{\mathcal{V}}_{t\text{-}1}^d(x_{t\text{-}1})\, dx_{t\text{-}1} \right]$$

- Inference in Gaussian Markov Chain!
- Special attention for $\beta(x_T)$, it includes the end-cost function:

$$q'_{\beta_T}(x_T) = \text{Proj}\left[q^{\setminus \beta_T}(x_T)\ \phi(x_T) \right].$$

EP for Graphs - Updating α and β

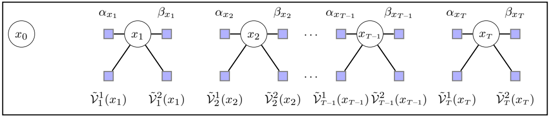


$$q'_{\alpha_t}(x_t) = \text{Proj}\left[q^{\backslash\alpha_t}(x_t) \int \mathcal{F}(x_{t\text{-}1}, x_t)\alpha(x_{t\text{-}1}) \prod_{d=1}^{D} \tilde{\mathcal{V}}_{t\text{-}1}^d(x_{t\text{-}1})\, dx_{t\text{-}1} \right]$$

- Inference in Gaussian Markov Chain!
- Special attention for $\beta(x_T)$, it includes the end-cost function:

$$q'_{\beta_T}(x_T) = \text{Proj}\left[q^{\backslash\beta_T}(x_T)\ \phi(x_T) \right].$$

EP for Graphs - Updating α and β

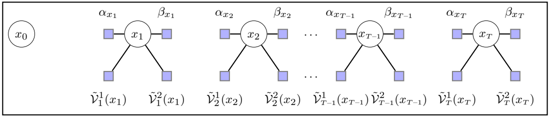


$$q'_{\alpha_t}(x_t) = \mathrm{Proj}\left[q^{\setminus\alpha_t}(x_t) \int \mathcal{F}(x_{t\text{-}1}, x_t)\alpha(x_{t\text{-}1}) \prod_{d=1}^{D} \tilde{\mathcal{V}}_{t\text{-}1}^d(x_{t\text{-}1})\, dx_{t\text{-}1}\right]$$

- Inference in Gaussian Markov Chain!
- Special attention for $\beta(x_T)$, it includes the end-cost function:

$$q'_{\beta_T}(x_T) = \mathrm{Proj}\left[q^{\setminus\beta_T}(x_T)\ \phi(x_T)\right].$$

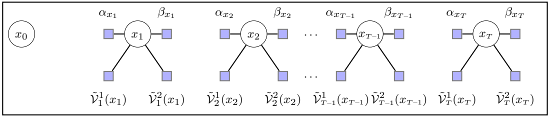
EP for Graphs - Updating Interval Constraints



$$q'_{\tilde{\mathcal{V}}_t^d}(x_t) = \mathrm{Proj}\left[q^{\setminus \tilde{\mathcal{V}}_t^d}(x_t)\ \mathcal{V}_t^d(x_t)\right]$$

- 1D case: moments of truncated Gaussian (Erf function)
- Multi-dimensional case
  - Rewrite to 1D prior and a conditional on other dimensions,
  - Use a 1D approximation for the prior.

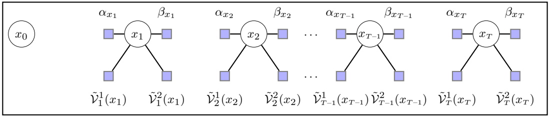
EP for Graphs - Updating Interval Constraints



$$q'_{\tilde{\mathcal{V}}_t^d}(x_t) = \mathrm{Proj}\left[q^{\setminus \tilde{\mathcal{V}}_t^d}(x_t)\ \mathcal{V}_t^d(x_t)\right]$$

- 1D case: moments of truncated Gaussian (Erf function)
- Multi-dimensional case
  - Rewrite to 1D prior and a conditional on other dimensions,
  - Use a 1D approximation for the prior.

EP for Graphs - Updating Interval Constraints

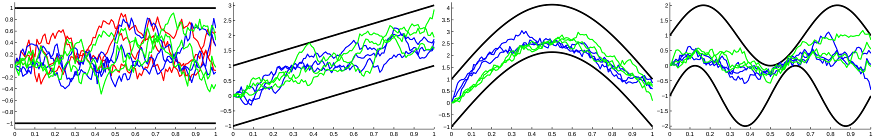


$$q'_{\tilde{\mathcal{V}}_t^d}(x_t) = \mathrm{Proj}\left[q^{\setminus \tilde{\mathcal{V}}_t^d}(x_t)\ \mathcal{V}_t^d(x_t) \right]$$

- 1D case: moments of truncated Gaussian (Erf function)
- Multi-dimensional case
  - Rewrite to 1D prior and a conditional on other dimensions,
  - Use a 1D approximation for the prior.

Experiments

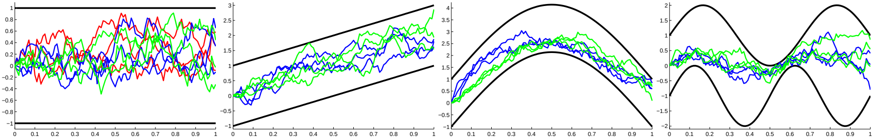
- Particle in a Box problem



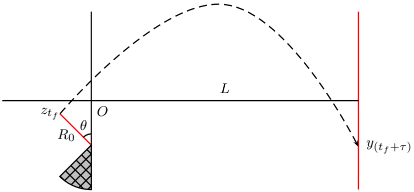
- Ball Throwing problem

Experiments

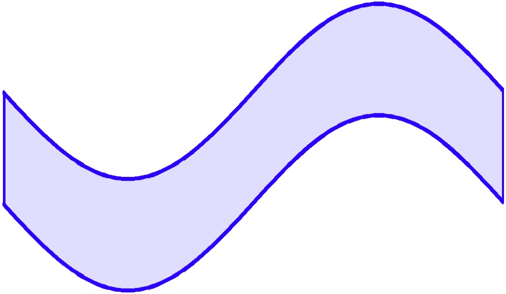
- Particle in a Box problem



- Ball Throwing problem

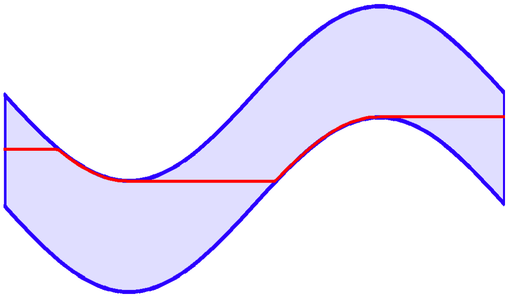


Particle in a box: Different Approximations



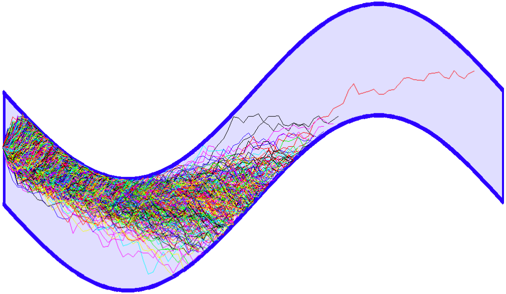
- Ignore noise and Variational Approximation
- MC Sampling
- Expectation Propagation

Particle in a box: Different Approximations



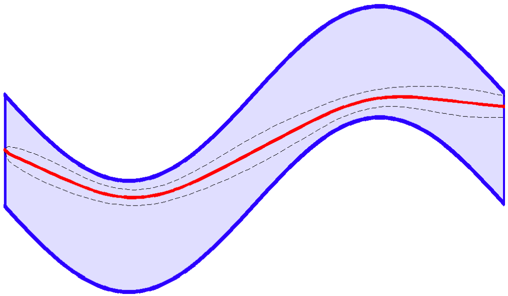
- Ignore noise and Variational Approximation
- MC Sampling
- Expectation Propagation

Particle in a box: Different Approximations



- Ignore noise and Variational Approximation
- MC Sampling
- Expectation Propagation

Particle in a box: Different Approximations



- Ignore noise and Variational Approximation
- MC Sampling
- Expectation Propagation

Particle in a box: Different Walls

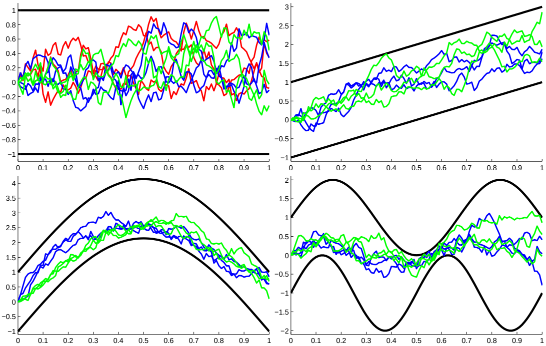
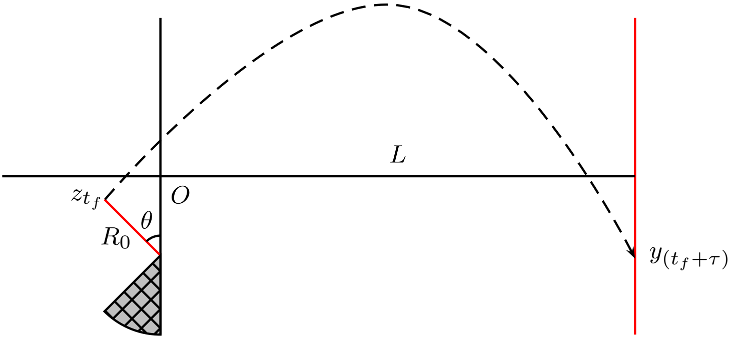


Table: Quantitative results of ‘Particle in a box’.

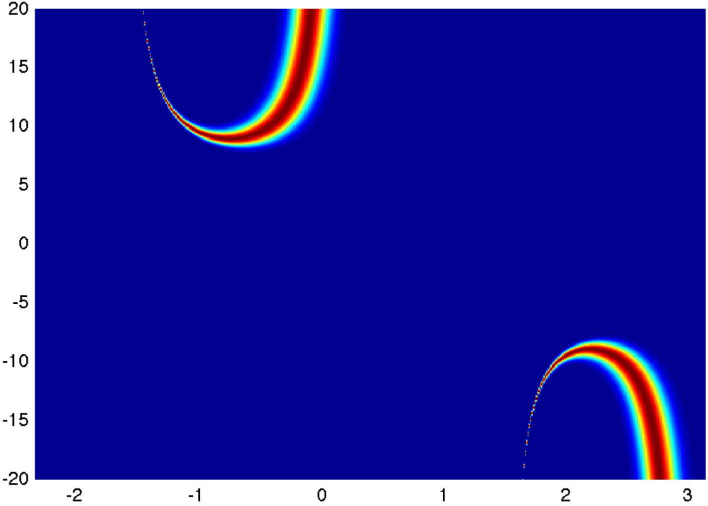
	EP		MC	
	Suc.	Cost	Suc.	Cost
Wall 1	24	1.04	23	0.70
Wall 2	24	1.83	23	1.47
Wall 3	22	5.02	13	4.35
Wall 4	15	2.50	7	1.98

Ball Throwing - Dynamics

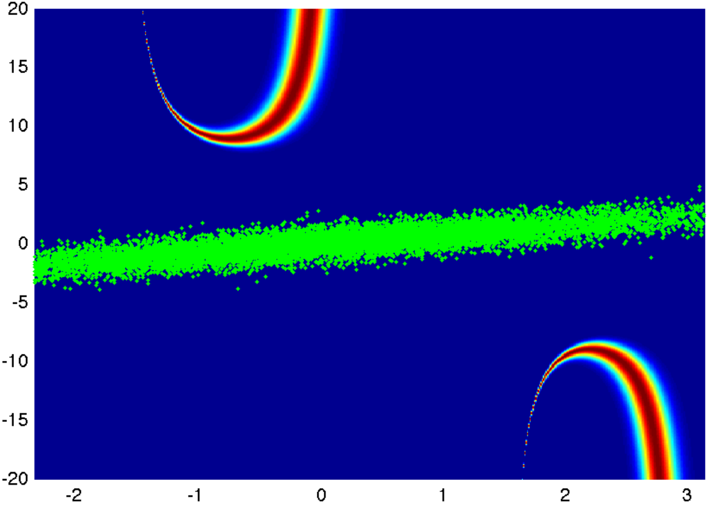


- Control speed of arm, not the direct position,
- Noise only between the force and the speed of the arm.
- End-cost function: $y(t_f + \tau)^2$.

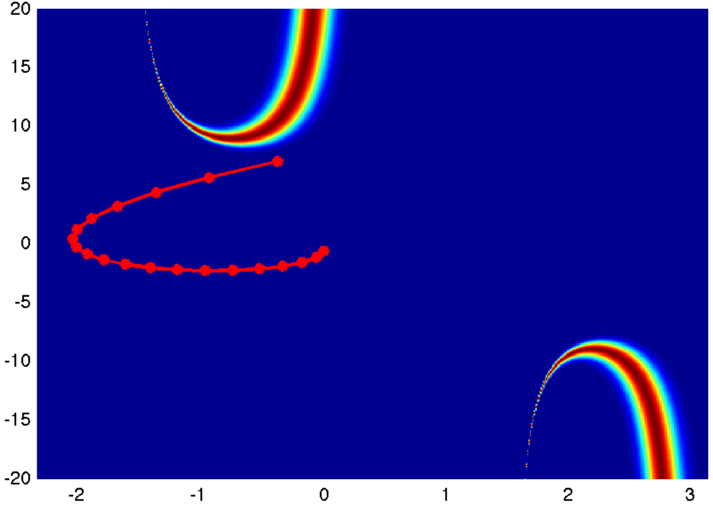
Ball Throwing - Goal Function - Sampling and EP



Ball Throwing - Goal Function - Sampling and EP



Ball Throwing - Goal Function - Sampling and EP

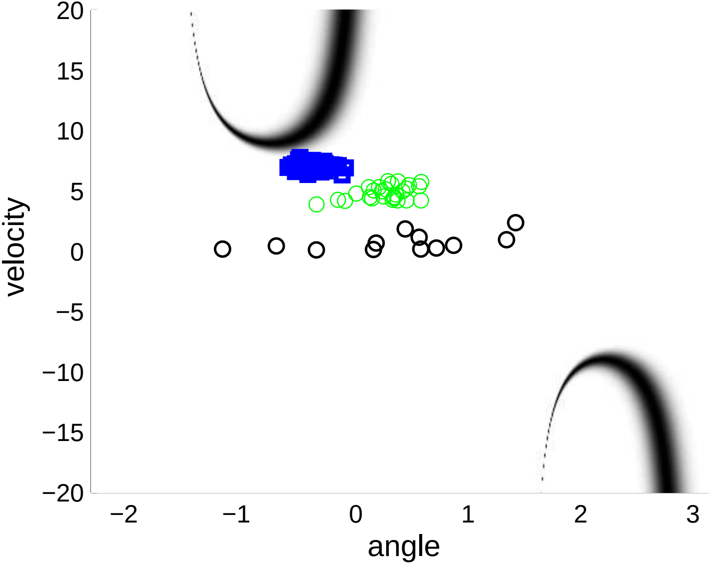


Ball Throwing - Quantitative Results

	Suc.	Acc.	Control	End	Total	Time
EP	50	50	1037	14	1050	1.0
MC 100k	33	13	540	219	759	0.9
MC 500k	39	27	614	188	803	3.4
MC 1000k	43	35	690	160	850	6.9

- Acceptable simulations (Acc.) $|y| \leq 75$.

Ball Throwing - Visual Results



Conclusions

- Stochastic Optimal Control with Hard Obstacles
- Using EP/MC approximations for the path integral

- EP outperforms MC sampler,
- Amount of necessary control w.r.t noise, important for success of MC.

- EP can be overly safe,
- But in complex domains obtains competitive overall cost.

Conclusions

- Stochastic Optimal Control with Hard Obstacles
- Using EP/MC approximations for the path integral

- EP outperforms MC sampler,
- Amount of necessary control w.r.t noise, important for success of MC.

- EP can be overly safe,
- But in complex domains obtains competitive overall cost.

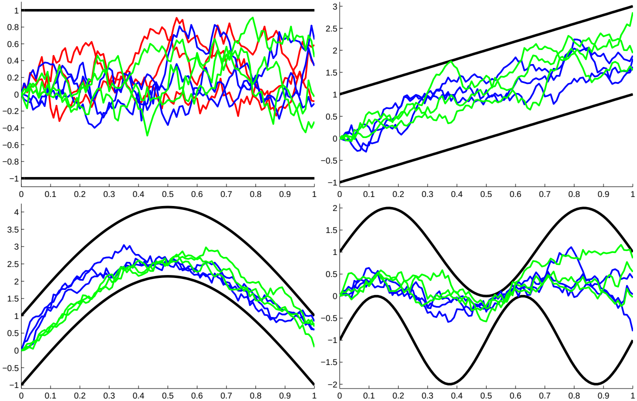
Conclusions

- Stochastic Optimal Control with Hard Obstacles
- Using EP/MC approximations for the path integral

- EP outperforms MC sampler,
- Amount of necessary control w.r.t noise, important for success of MC.

- EP can be overly safe,
- But in complex domains obtains competitive overall cost.

Questions?



1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
3
2.5
2
1.5
0.5
−0.5
4
3.5
−1.5
−2