Probabilistic Inference of Simulation Parameters via Parallel Differentiable Simulation

Eric Heiden[1] Christopher E. Denniston[1] David Millard[1] Fabio Ramos[2,3] Gaurav S. Sukhatme[1]

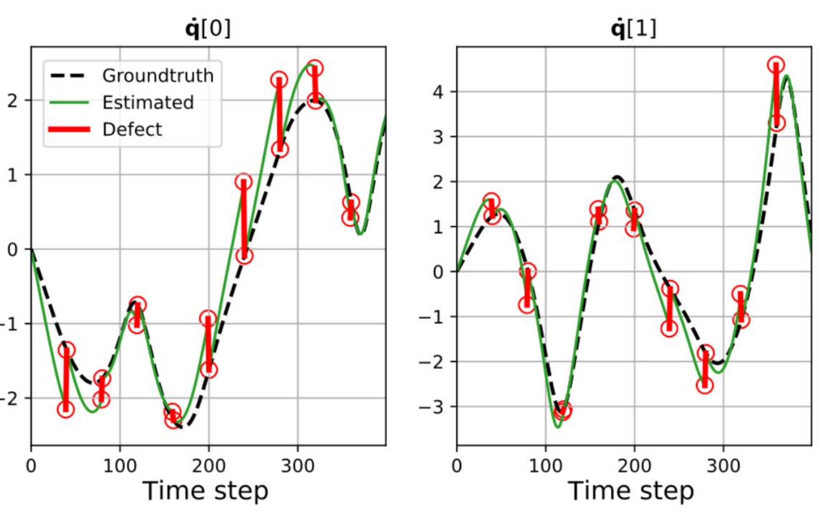
[1]University of Southern California, USA [2]Nvidia, USA [3]University of Sydney, Australia

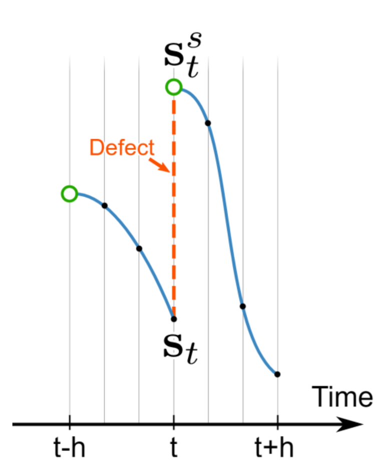
Why probabilistic parameter estimation?

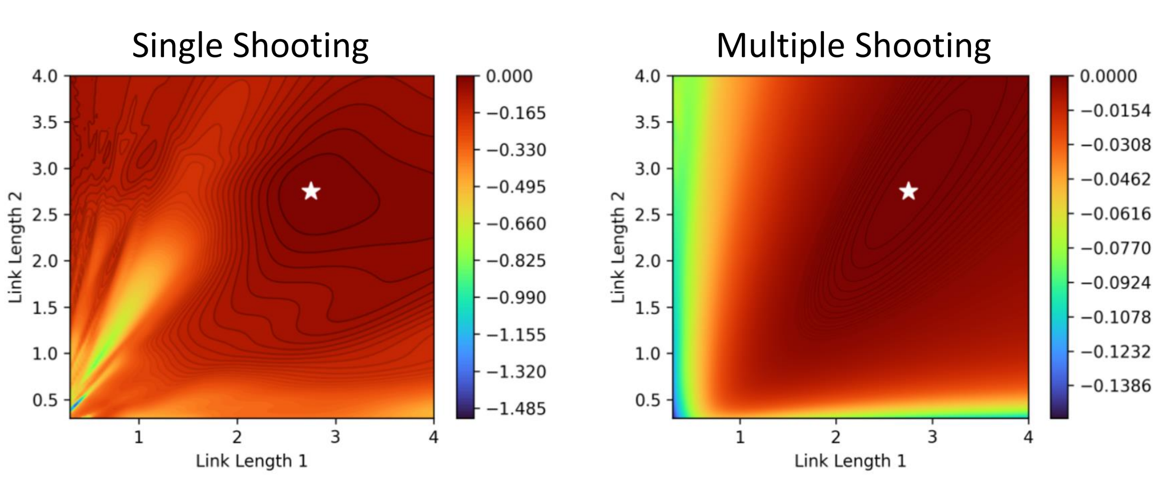
- Reality gap in simulators often due to incorrect simulation settings
- Inference of simulation parameters is difficult for chaotic, nonlinear systems
- Observations are noisy, state estimates are uncertain
- There can be multiple possible parameter settings explaining the same observed behavior

Smoother gradients with Multiple Shooting

- Split up trajectory into *shooting windows*
- Jointly estimate start states of shooting windows with simulation parameters
- Impose *defect constraints* to ensure continuity
- Yields smoother likelihood landscape than single-shooting







Particle-based Bayesian inference

Infer posterior $p(\theta|D_\mathcal{X})$ over simulation parameters $\theta \in \mathbb{R}^M$ and a set of trajectories $D_\mathcal{X}$ via Bayes' rule:

$$p(\theta|D_\mathcal{X}) \propto p(D_\mathcal{X}|\theta)\, p(\theta)$$

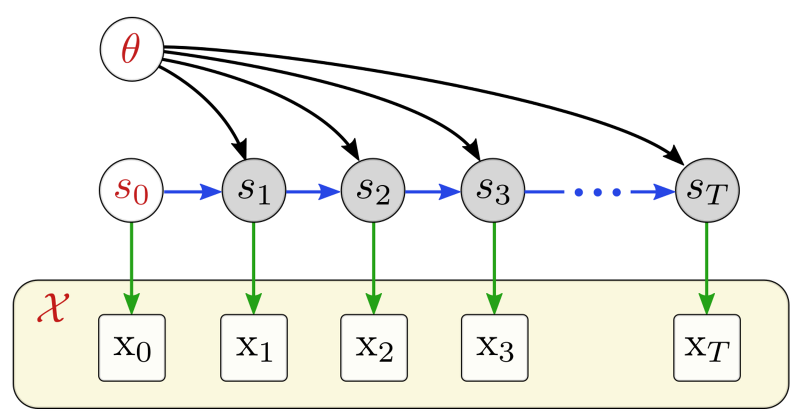
Hidden Markov Model:

initial state s_0, observation function f_obs , simulation function f_sim

$$\mathcal{X} = f_\text{obs}([s]_{t=1}^T)$$

$$f_\text{sim}(\theta, s_0) = [s]_{t=1}^T$$

$$D_\mathcal{X}^\text{sim} = \left[f_\text{obs}\left(f_\text{sim}(\theta, s_0^\text{real})\right)\right]$$



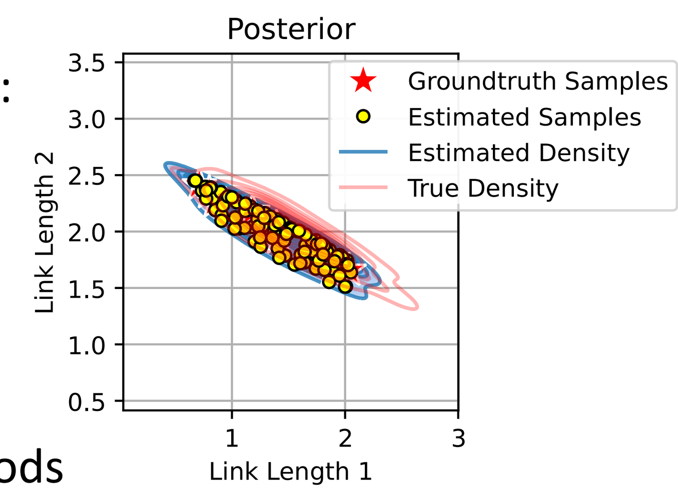
Objective: minimize KL divergence

$$d_\text{KL}\left[p(D_\mathcal{X}^\text{sim}|\theta^\text{sim})\, p(\theta^\text{sim}) \parallel p(D_\mathcal{X}^\text{real}|\theta^\text{real})\, p(\theta^\text{real})\right]$$

simulation reality

Stein Variational Gradient Descent (SVGD):

- represents posterior distribution by particles
- can accurately estimate complex multimodal distributions
- leverages parallel, differentiable likelihoods



Our method

Constrained Stein Variational Gradient Descent (SVGD):

Extends SVGD to include hard constraints in the particle estimation:

- Parameter limit constraints
- Defect constraints to realize Multiple Shooting approach

$$\dot\theta = \phi(\theta) \quad -\lambda\frac{\partial g(\theta)}{\partial\theta} - c g(\theta)\frac{\partial g(\theta)}{\partial\theta}, \qquad \dot\lambda = g(\theta) \qquad \text{CSVGD}$$

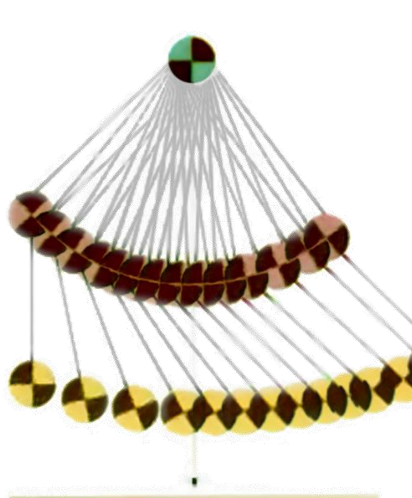
$$\phi(\cdot) = \frac{1}{N}\sum_{j=1}^N \left[k(\theta_j, \theta)\nabla_{\theta_j}\log p(D_\mathcal{X}|\theta) + \nabla_{\theta_j}k(\theta_j,\theta)\right] \qquad \text{SVGD}$$

where $k(\cdot,\cdot)$ is a positive definite kernel (RBF), N is the number of particles

Experiments

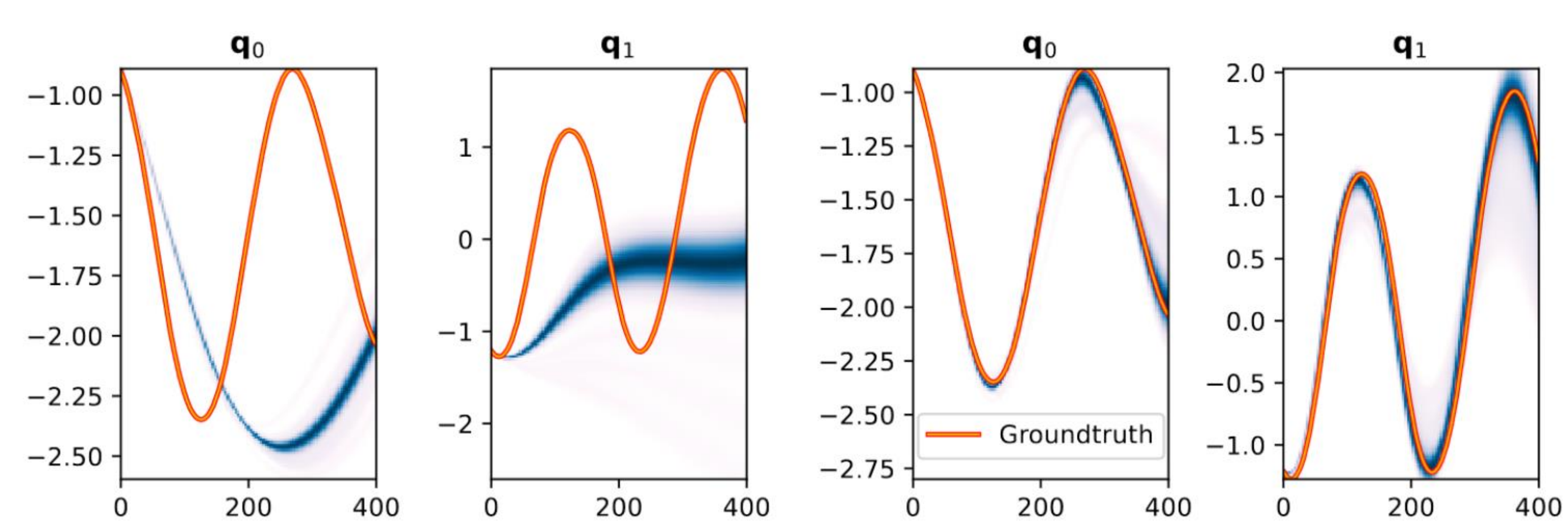
Double pendulum

Infer 11 parameters of a real double pendulum system



Link	Parameter	Minimum	Maximum
Link 1	Mass	0.05 kg	0.5 kg
	I_{xx}	0.002 kg m^2	1.0 kg m^2
	COM x	−0.2 m	0.2 m
	COM y	−0.2 m	0.2 m
	Joint friction	0.0	0.5
Link 2	Length	0.08 m	0.3 m
	Mass	0.05 kg	0.5 kg
	I_{xx}	0.002 kg m^2	1.0 kg m^2
	COM x	−0.2 m	0.2 m
	COM y	−0.2 m	0.2 m
	Joint friction	0.0	0.5

Trajectory densities resulting from the estimated parameter distributions:

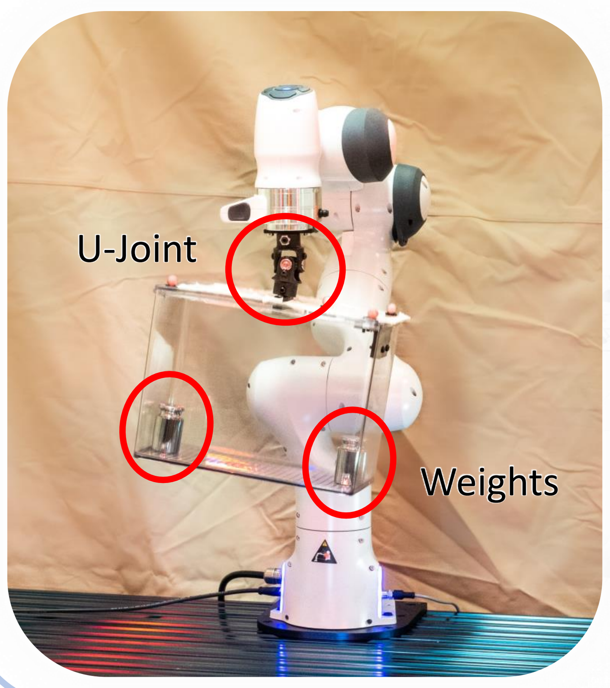


SVGD CSVGD

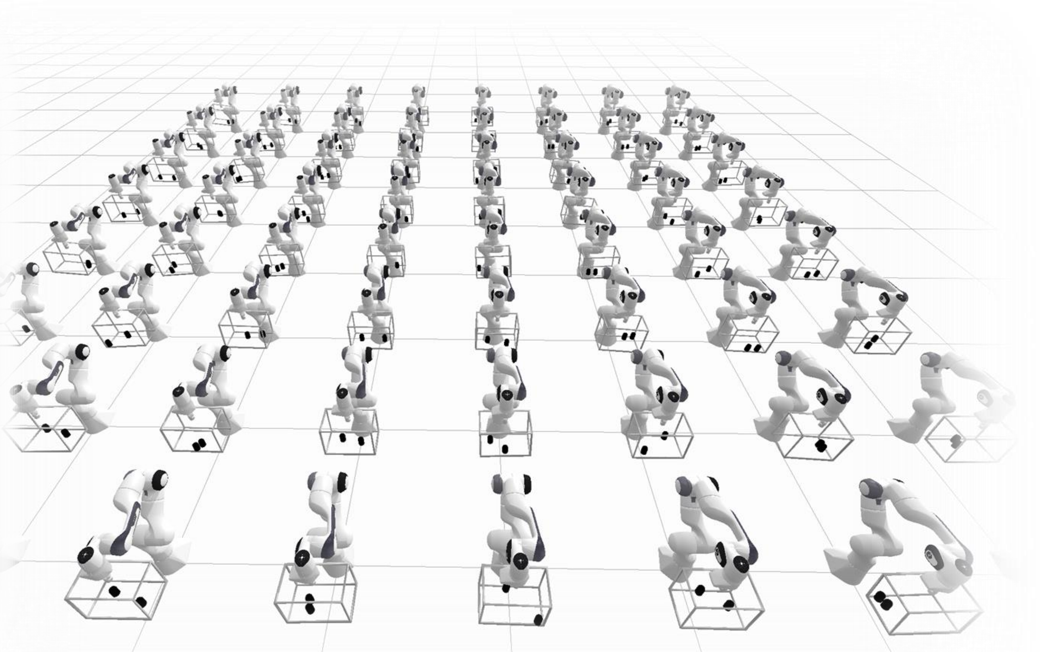
Underactuated mechanism

Infer the locations of the two 500g weights in the box attached to a Panda robot arm through a universal joint

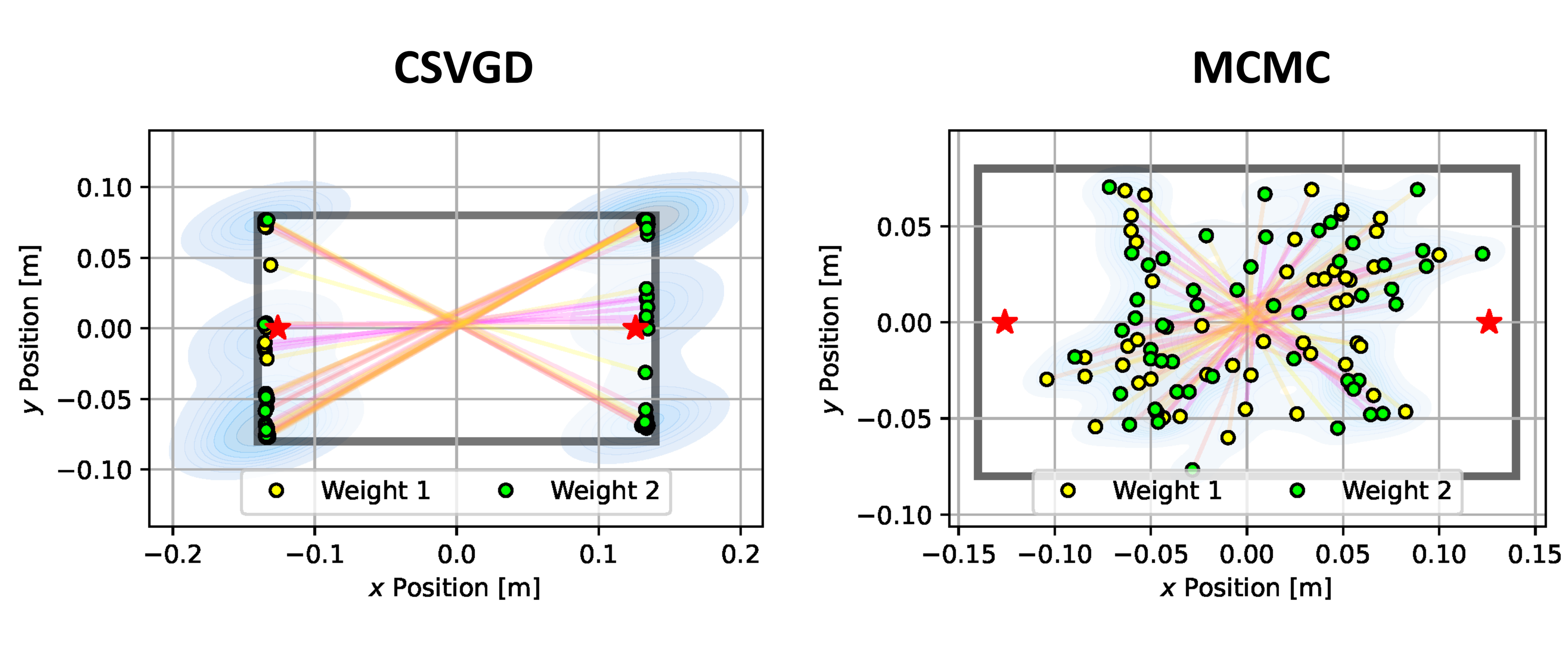
Setup on real robot



GPU-based, parallel differentiable simulations



Inferred posterior distributions:



ICRA 2022 IEEE International Conference on Robotics and Automation
ROBOTIC EMBEDDED SYSTEMS LABORATORY
USC University of Southern California
NVIDIA