Probabilistic inverse optimal control for non-linear partially observable systems disentangles perceptual uncertainty and behavioral costs

Dominik Straub*, Matthias Schultheis*, Heinz Koeppl, Constantin A. Rothkopf
Centre for Cognitive Science
Technische Universität Darmstadt, Germany

*equal contribution

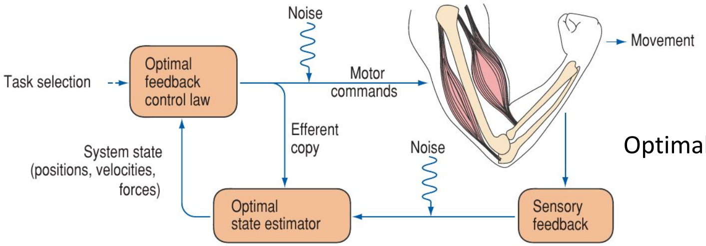
TECHNISCHE UNIVERSITÄT DARMSTADT

Centre for Cognitive Science

NeurIPS 2023

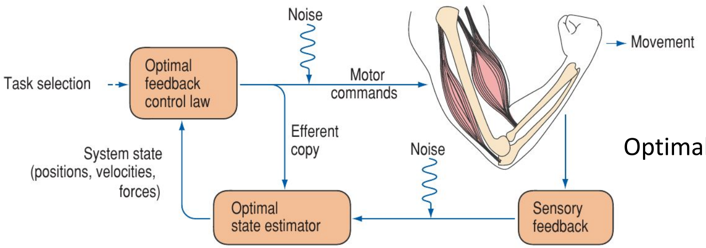
NEURAL INFORMATION PROCESSING SYSTEMS

Goal: understanding sensorimotor behavior



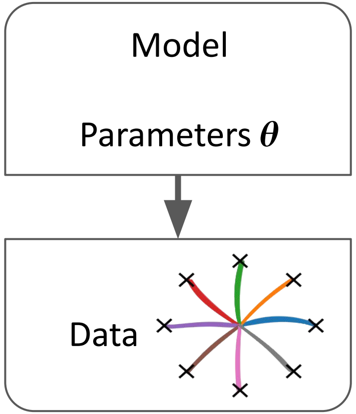
Optimal feedback control (Scott, 2008)

Goal: understanding sensorimotor behavior



Optimal feedback control (Scott, 2008)

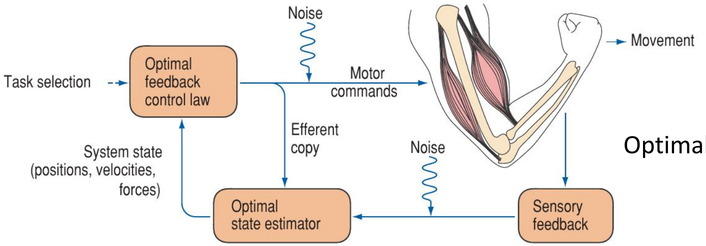
Optimal control



find control law to **maximize expected reward**

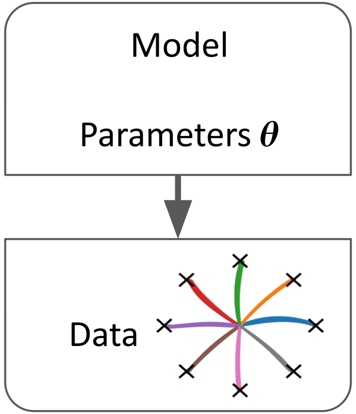
Dynamics may be **non-linear**,
costs **non-quadratic,**
partial and noisy observations,
non-Gaussian noise.

Goal: understanding sensorimotor behavior



Optimal feedback control (Scott, 2008)

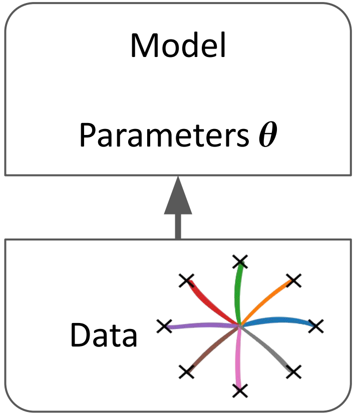
Optimal control



find control law to **maximize expected reward**

Dynamics may be **non-linear**, costs **non-quadratic,** **partial and noisy observations,** **non-Gaussian** noise.

Inverse optimal control



infer parameters of agent's **internal model** and **cost function**

Agent's **internal beliefs and control signals** are **unobserved.**

Related work

Classic work on inferring agent's cost function in psychology & economics

Mosteller, F. and Nogee, P. (1951)
Kahneman, D. and Tversky, A. (1979)
Körding, K. P. and Wolpert, D. M. (2004)

Inverse Optimal Control (IOC)

Ziebart, B. D., Maas, A. L., Bagnell, J. A., and Dey, A. K. (2008)
Levine, S. and Koltun, V. (2012)
Garg, D., Chakraborty, S., Cundy, C., Song, J., and Ermon, S. (2021)

Partially-observable IOC

Choi, J. D., and Kim, K. E. (2011)
Bogert, K., Lin, J. F. S., Doshi, P., and Kulic, D. (2016)
Kwon, M., Daptardar, S., Schrater, P. R., and Pitkow, X. (2020)
Schultheis, M., Straub, D., and Rothkopf, C. A. (2021)

Related work

Classic work on inferring agent's cost function in psychology & economics

Mosteller, F. and Nogee, P. (1951)
Kahneman, D. and Tversky, A. (1979)
Körding, K. P. and Wolpert, D. M. (2004)

Inverse Optimal Control (IOC)

Ziebart, B. D., Maas, A. L., Bagnell, J. A., and Dey, A. K. (2008)
Levine, S. and Koltun, V. (2012)
Garg, D., Chakraborty, S., Cundy, C., Song, J., and Ermon, S. (2021)

Partially-observable IOC

Choi, J. D., and Kim, K. E. (2011)
Bogert, K., Lin, J. F. S., Doshi, P., and Kulic, D. (2016)
Kwon, M., Daptardar, S., Schrater, P. R., and Pitkow, X. (2020)
Schultheis, M., Straub, D., and Rothkopf, C. A. (2021)

We present the first method for
continuous,
partially observable,
non-linear,
stochastic
systems with **unobserved control signals**

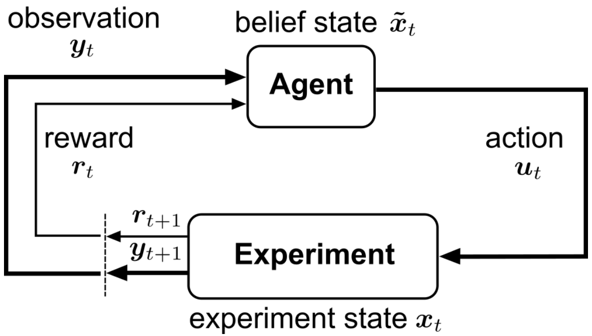
Optimal control: problem formulation

Dynamical system

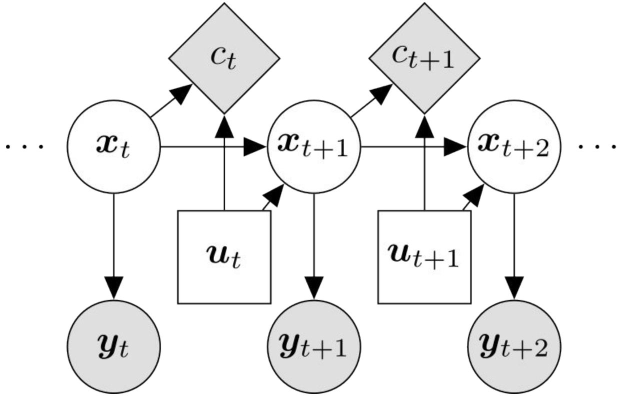
$$\boldsymbol{x}_{t+1} = f(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{v}_t) \qquad \boldsymbol{v}_t \sim \mathcal{N}(0, I)$$

Observation model

$$\boldsymbol{y}_t = h(\boldsymbol{x}_t, \boldsymbol{w}_t) \qquad \boldsymbol{w}_t \sim \mathcal{N}(0, I)$$



Agent-experiment loop (Schultheis et al., 2021)



Optimal control: problem formulation

Dynamical system

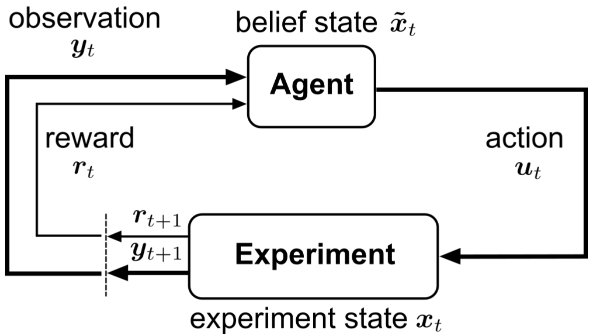
$$\boldsymbol{x}_{t+1} = f(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{v}_t) \qquad \boldsymbol{v}_t \sim \mathcal{N}(0, I)$$

Observation model

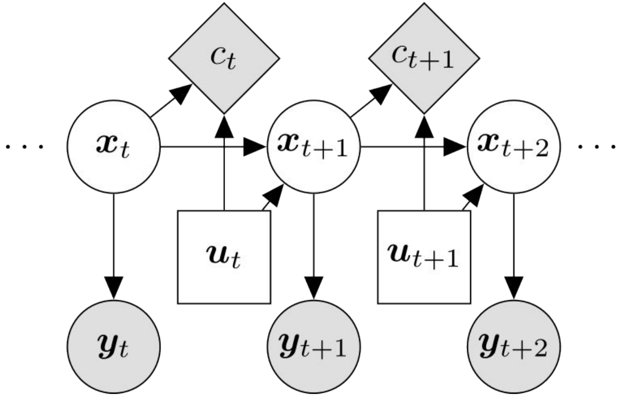
$$\boldsymbol{y}_t = h(\boldsymbol{x}_t, \boldsymbol{w}_t) \qquad \boldsymbol{w}_t \sim \mathcal{N}(0, I)$$

Cost function

$$J = \mathbb{E}\left[c_T(\boldsymbol{x}_T) + \sum_{t=1}^{T-1} c_t(\boldsymbol{x}_t, \boldsymbol{u}_t)\right]$$



Agent-experiment loop (Schultheis et al., 2021)



Optimal control: problem formulation

Dynamical system

$$\boldsymbol{x}_{t+1} = f(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{v}_t) \qquad \boldsymbol{v}_t \sim \mathcal{N}(0, I)$$

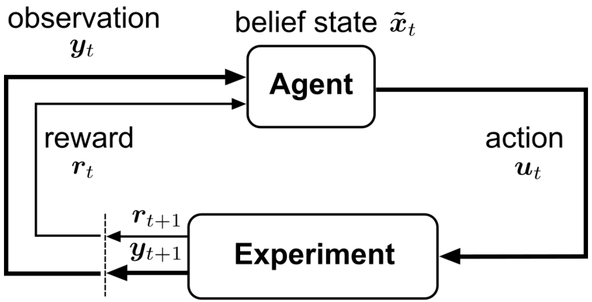
Observation model

$$\boldsymbol{y}_t = h(\boldsymbol{x}_t, \boldsymbol{w}_t) \qquad \boldsymbol{w}_t \sim \mathcal{N}(0, I)$$

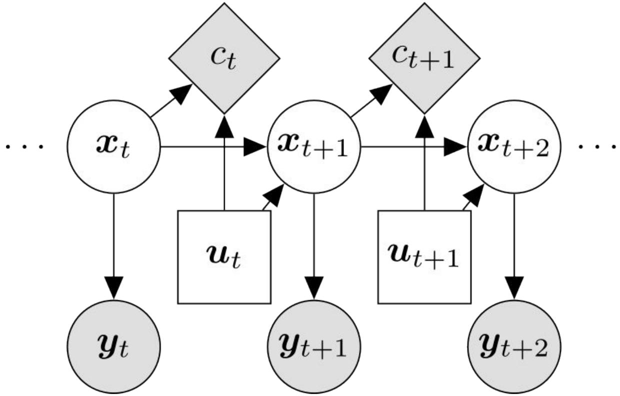
Cost function

$$J = \mathbb{E}\left[c_T(\boldsymbol{x}_T) + \sum_{t=1}^{T-1} c_t(\boldsymbol{x}_t, \boldsymbol{u}_t)\right]$$

Approximately optimal solution: iLQG (Li & Todorov, 2007)

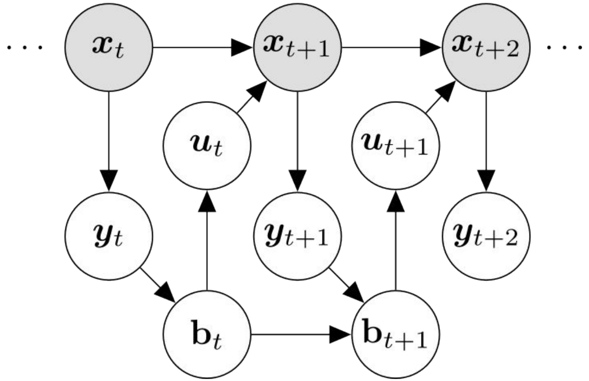


Agent-experiment loop (Schultheis et al., 2021)



Probabilistic inverse optimal control

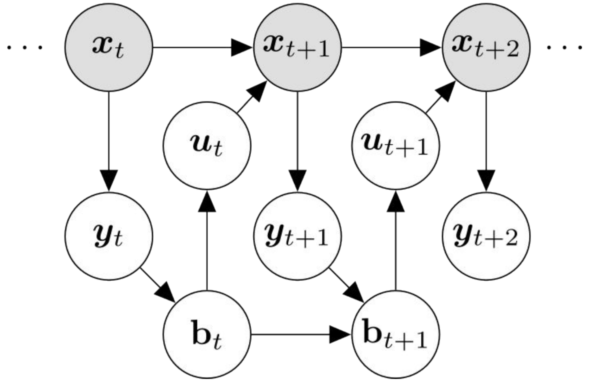
Goal: maximize likelihood $p(\boldsymbol{x}_{1:T} \mid \boldsymbol{\theta}) = p(\boldsymbol{x}_1 \mid \boldsymbol{\theta}) \prod_{t=1}^{T-1} p(\boldsymbol{x}_{t+1} \mid \boldsymbol{x}_{1:t}, \boldsymbol{\theta})$



Probabilistic inverse optimal control

Goal: maximize likelihood $p(\boldsymbol{x}_{1:T} \mid \boldsymbol{\theta}) = p(\boldsymbol{x}_1 \mid \boldsymbol{\theta}) \prod_{t=1}^{T-1} p(\boldsymbol{x}_{t+1} \mid \boldsymbol{x}_{1:t}, \boldsymbol{\theta})$

Problem: Each state depends on all previous states
(via agent's belief, which is unobservable).

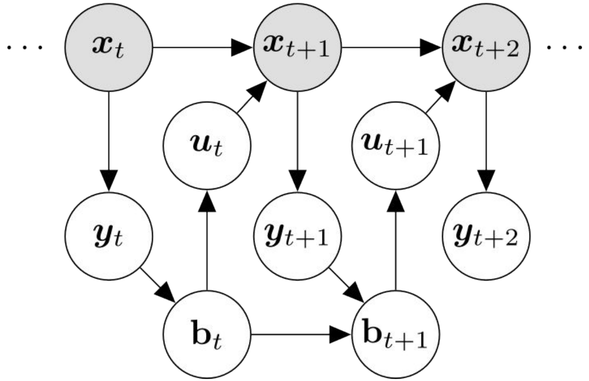


Probabilistic inverse optimal control

Goal: maximize likelihood $p(\boldsymbol{x}_{1:T} \mid \boldsymbol{\theta}) = p(\boldsymbol{x}_1 \mid \boldsymbol{\theta}) \prod_{t=1}^{T-1} p(\boldsymbol{x}_{t+1} \mid \boldsymbol{x}_{1:t}, \boldsymbol{\theta})$

Problem: Each state depends on all previous states
(via agent's belief, which is unobservable).

Solution: two central ideas



Probabilistic inverse optimal control

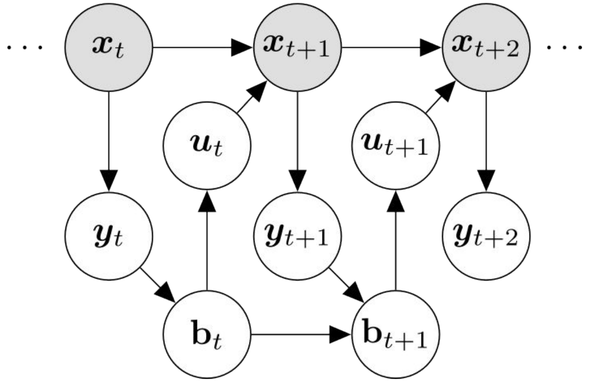
Goal: maximize likelihood $p(\boldsymbol{x}_{1:T} \mid \boldsymbol{\theta}) = p(\boldsymbol{x}_1 \mid \boldsymbol{\theta}) \prod_{t=1}^{T-1} p(\boldsymbol{x}_{t+1} \mid \boldsymbol{x}_{1:t}, \boldsymbol{\theta})$

Problem: Each state depends on all previous states (via agent's belief, which is unobservable).

Solution: two central ideas

1. Joint system of states and beliefs is Markovian.

$$\begin{bmatrix} \boldsymbol{x}_{t+1} \\ \mathbf{b}_{t+1} \end{bmatrix} \sim p(\boldsymbol{x}_{t+1}, \mathbf{b}_{t+1} \mid \boldsymbol{x}_t, \mathbf{b}_t)$$



Probabilistic inverse optimal control

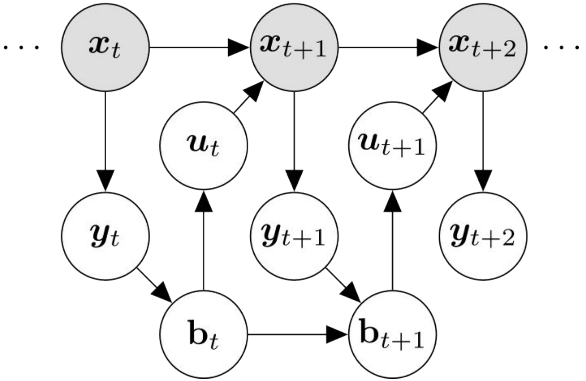
Goal: maximize likelihood $p(\boldsymbol{x}_{1:T} \mid \boldsymbol{\theta}) = p(\boldsymbol{x}_1 \mid \boldsymbol{\theta}) \prod_{t=1}^{T-1} p(\boldsymbol{x}_{t+1} \mid \boldsymbol{x}_{1:t}, \boldsymbol{\theta})$

Problem: Each state depends on all previous states
(via agent's belief, which is unobservable).

Solution: two central ideas

1. Joint system of states and beliefs is Markovian.

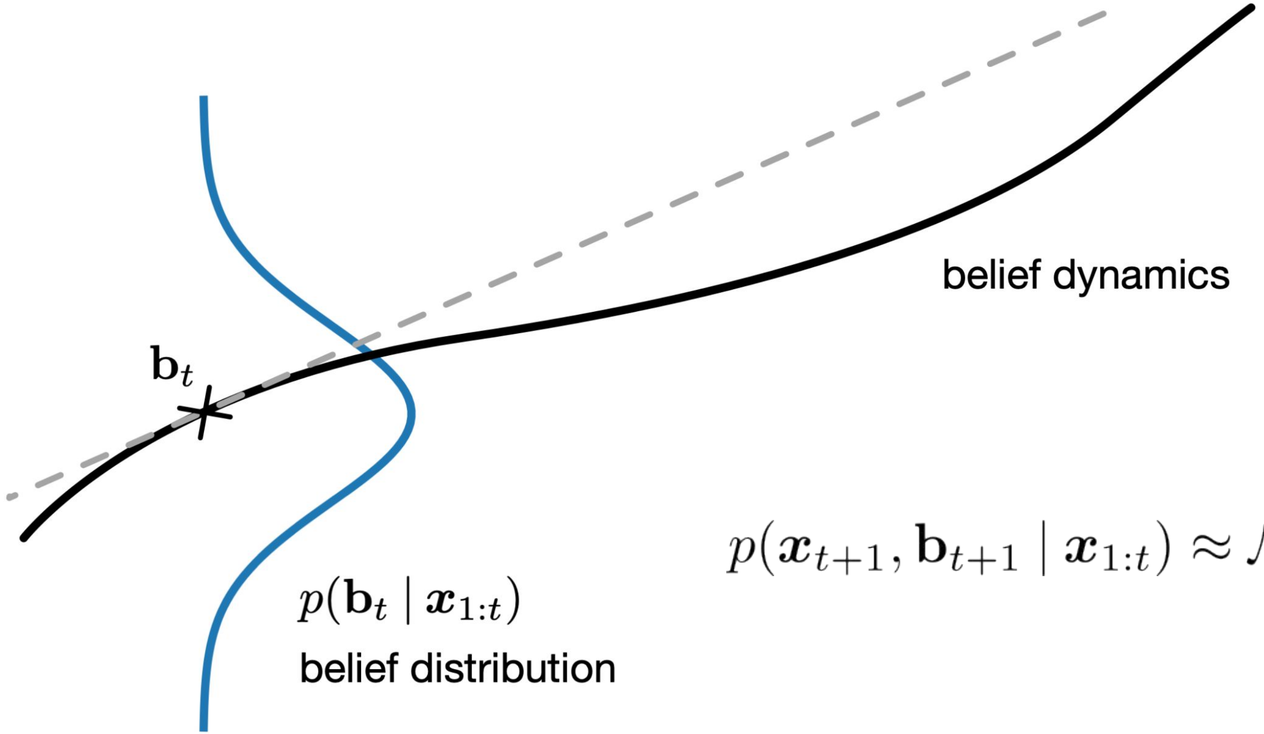
$$\begin{bmatrix} \boldsymbol{x}_{t+1} \\ \mathbf{b}_{t+1} \end{bmatrix} \sim p(\boldsymbol{x}_{t+1}, \mathbf{b}_{t+1} \mid \boldsymbol{x}_t, \mathbf{b}_t)$$



2. Belief tracking: distribution of agent's belief given past states

$$p(\mathbf{b}_t \mid \boldsymbol{x}_{1:t})$$

Tractable likelihood computation

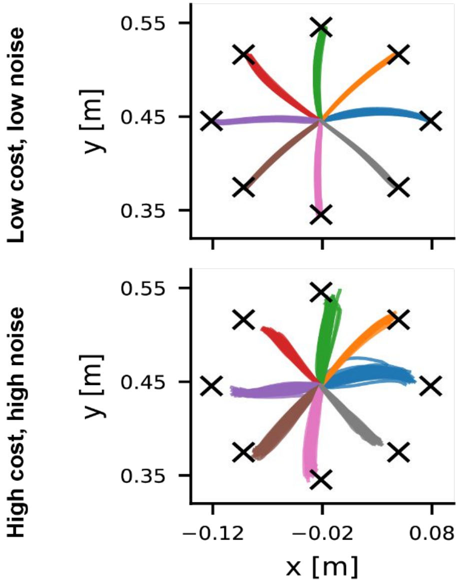


$$p(\boldsymbol{x}_{t+1}, \mathbf{b}_{t+1} \mid \boldsymbol{x}_{1:t}) \approx \mathcal{N}(\mu_t, \Sigma_t)$$

Results: non-linear reaching task

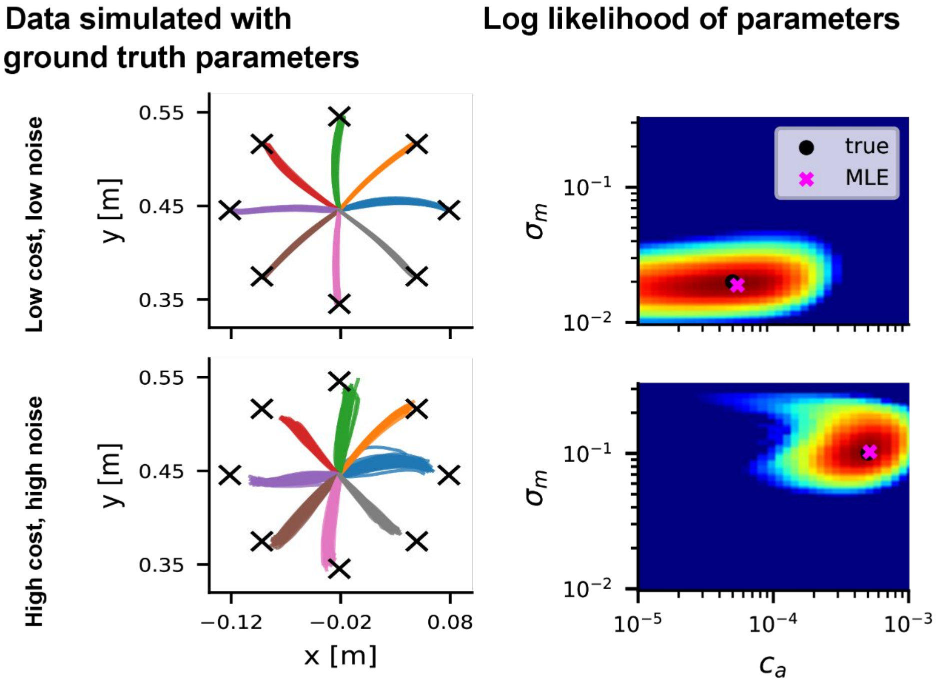
non-linear, partially observable reaching task (Li & Todorov, 2007)

Data simulated with ground truth parameters



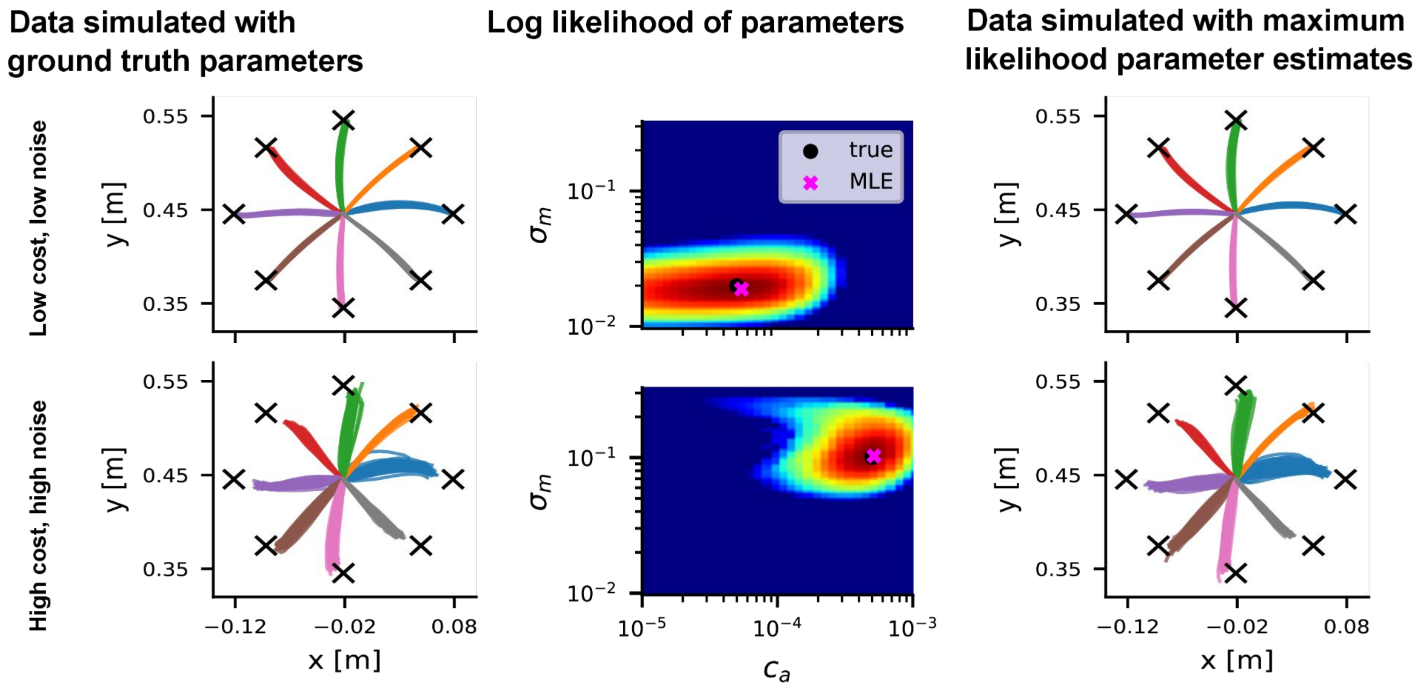
Results: non-linear reaching task

non-linear, partially observable reaching task (Li & Todorov, 2007)

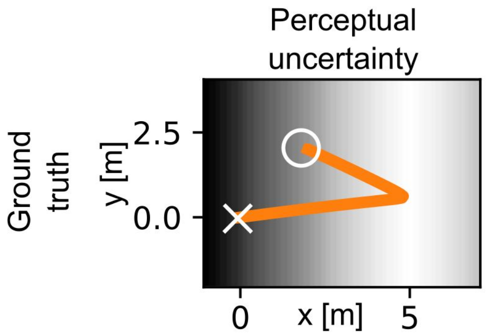


Results: non-linear reaching task

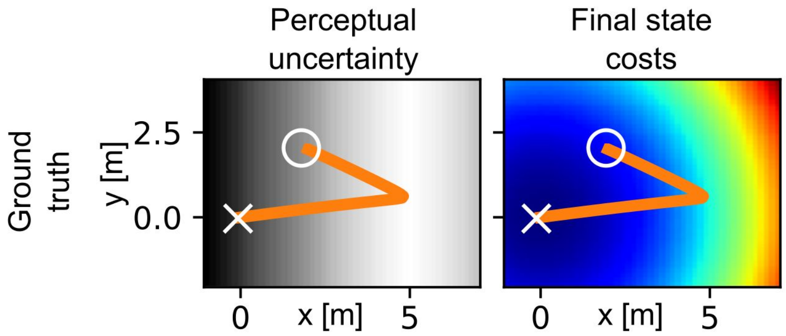
non-linear, partially observable reaching task (Li & Todorov, 2007)



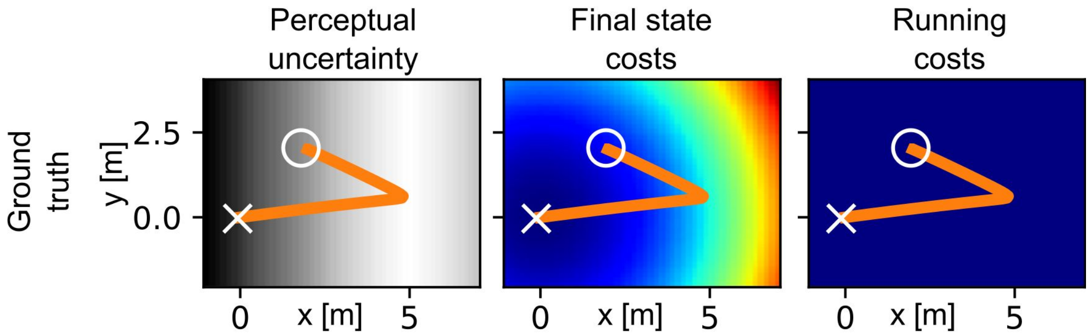
Information-seeking behavior in the light-dark domain



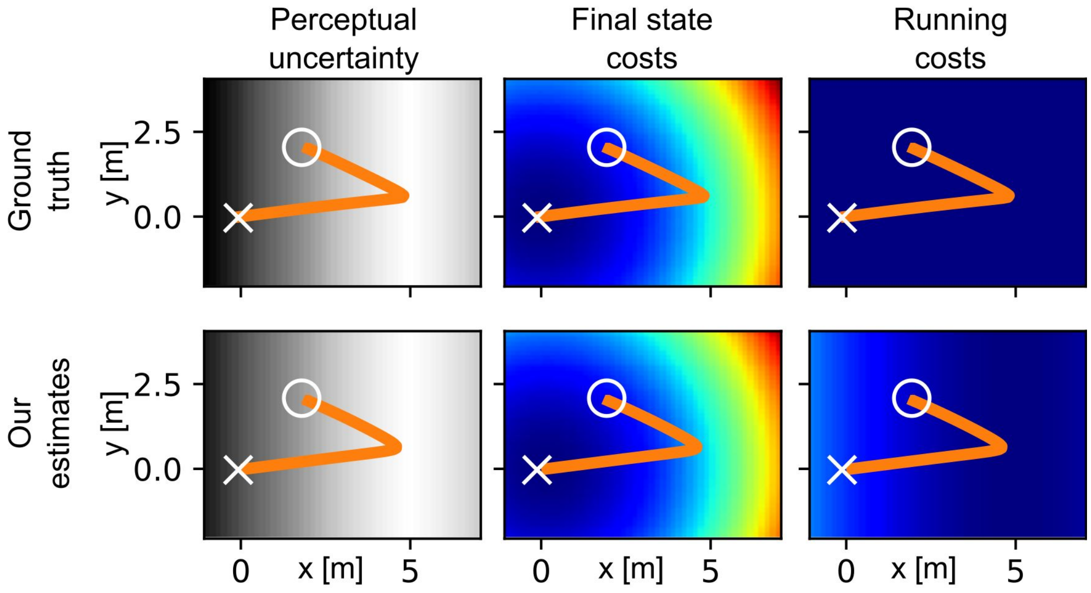
Information-seeking behavior in the light-dark domain



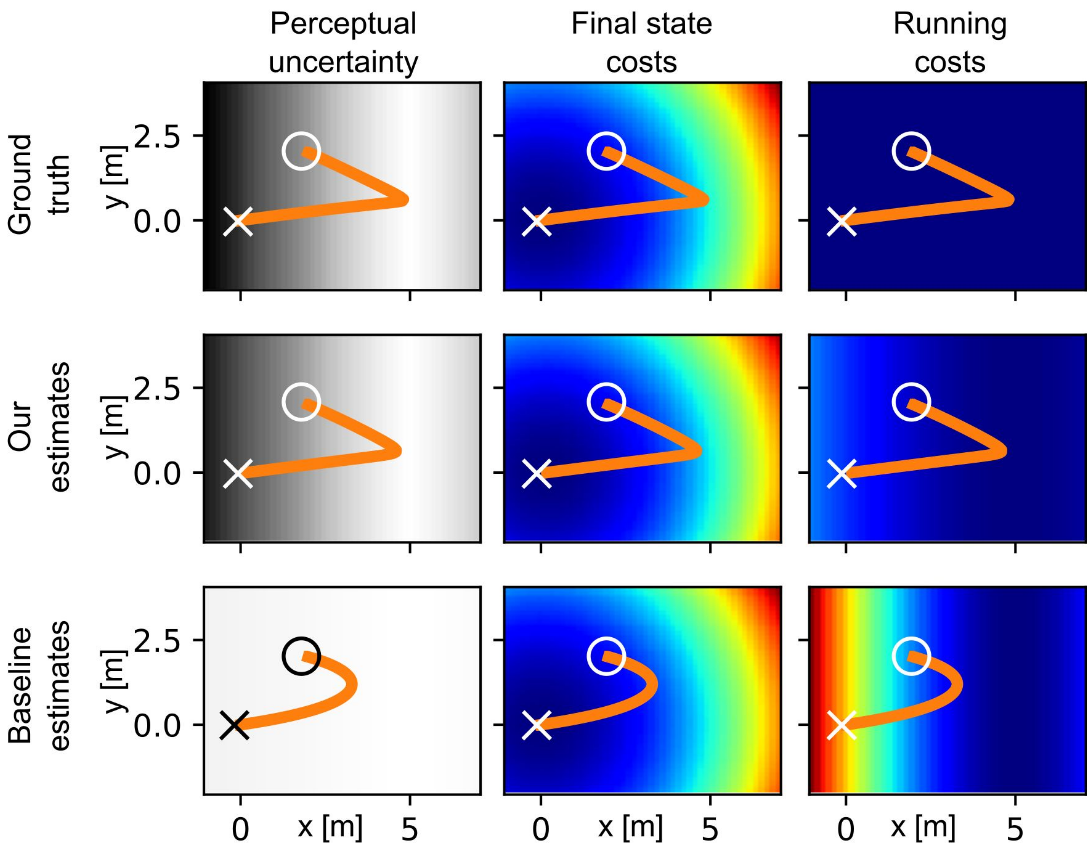
Information-seeking behavior in the light-dark domain



Information-seeking behavior in the light-dark domain



Information-seeking behavior in the light-dark domain



Conclusion

Probabilistic inverse optimal control method for
non-linear
partially observable
systems with **unobserved control signals**

Efficient likelihood maximization by
first-order Taylor approximation
leveraging data for linearization points

Applicable to cognitive science and neuroscience
infers cost function and sensorimotor noise characteristics
can be used to disentangle perceptual uncertainty and behavioral costs

TECHNISCHE UNIVERSITÄT DARMSTADT

Centre for Cognitive Science

Code available
https://github.com/RothkopfLab/nioc-neurips

Conclusion

Probabilistic inverse optimal control method for
- **non-linear**
- **partially observable**

systems with **unobserved control signals**

Efficient likelihood maximization by
- first-order Taylor approximation
- leveraging data for linearization points

Applicable to cognitive science and neuroscience
- infers cost function and sensorimotor noise characteristics
- can be used to disentangle perceptual uncertainty and behavioral costs

Check out our paper and code and come by our poster!

TECHNISCHE UNIVERSITÄT DARMSTADT

Centre for Cognitive Science

Code available
https://github.com/RothkopfLab/nioc-neurips