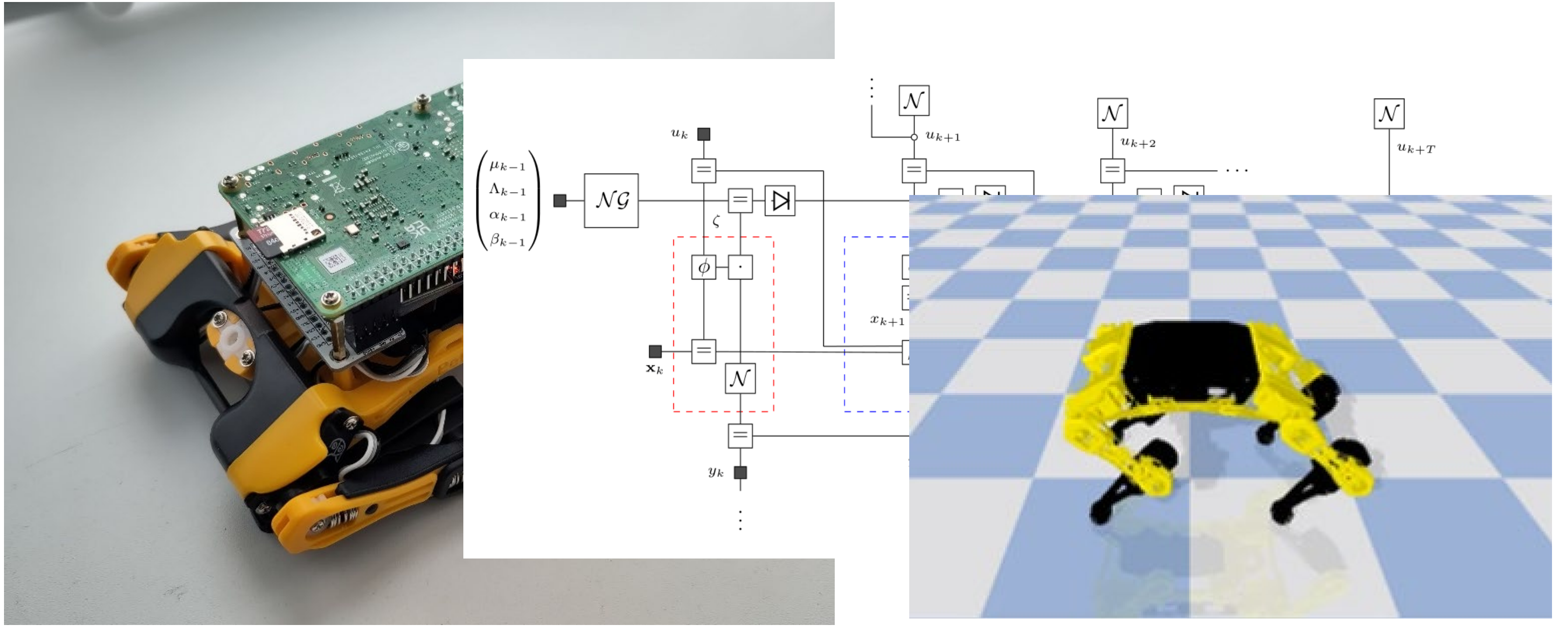
Planning to avoid ambiguous states through Gaussian approximations to non-linear sensors

Wouter M. Kouw
International Workshop on Active Inference 2024

EINDHOVEN UNIVERSITY OF TECHNOLOGY
BIASlab

Background / context



I am interested in discrete-time active inference with continuous states and actions.

Background / context

- What is the simplest model structure that generates explorative / exploitative behaviour?
  - A linear Gaussian state-space model?

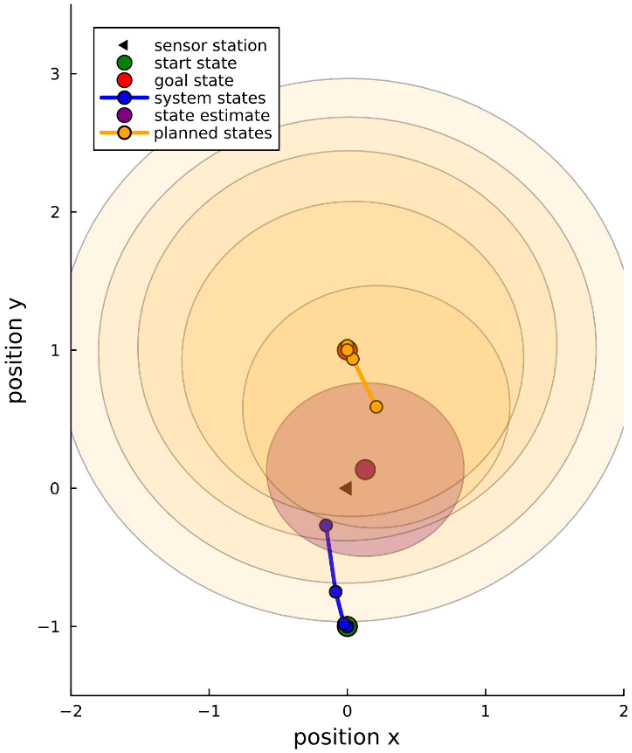
$$p(x_k|x_{k-1}, u_k) = \mathcal{N}(x_k|Ax_{k-1} + Bu_k, Q)$$
$$p(y_k|x_k) = \mathcal{N}(y_k|Cx_k, R)$$

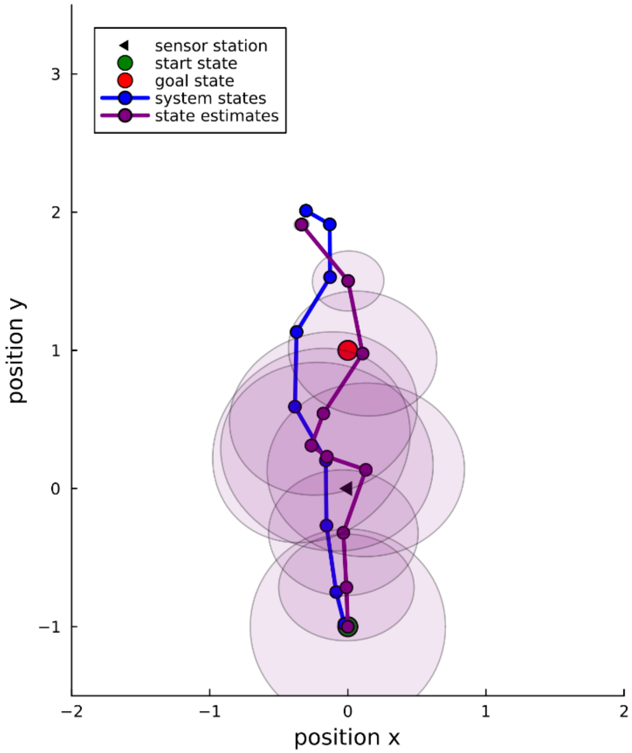
- No, because ambiguity will be constant over states (Koudahl, Kouw & de Vries, 2021);

$$\text{Ambiguity} = \frac{1}{2}\left(\mathrm{D_y} \ln 2\pi e + \ln|R|\right)$$

- What if you had a nonlinear observation function?

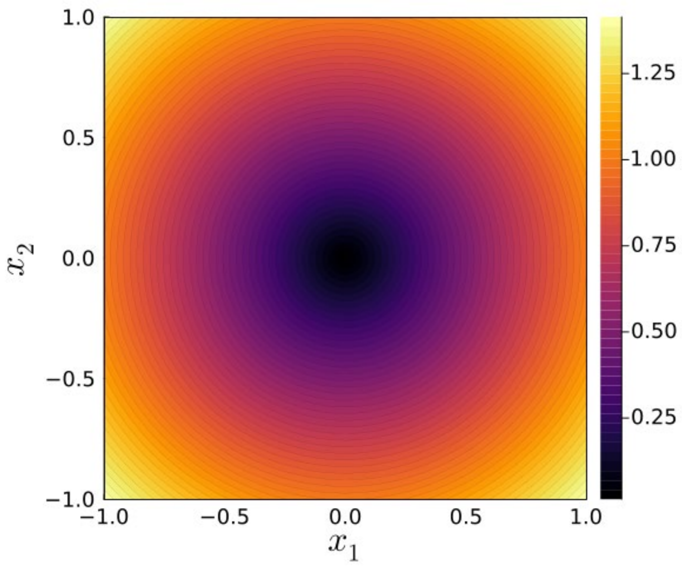
Challenge

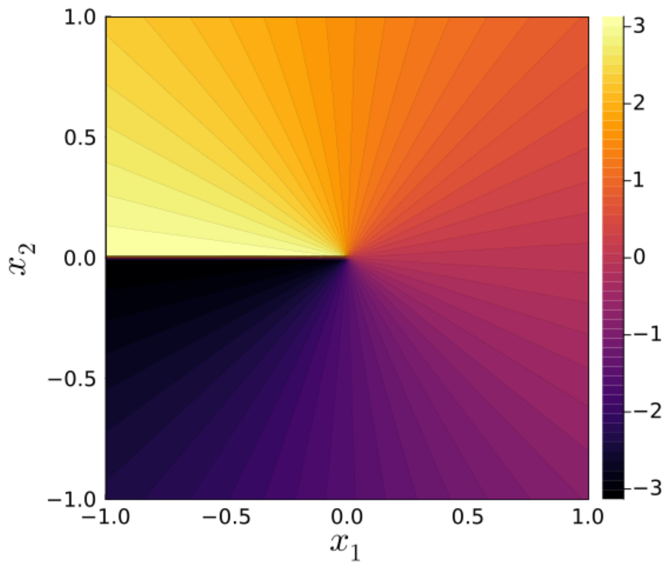

sensor station
start state
goal state
system states
state estimate
planned states
position y
position x
3
2
1
0
−1
−2
−1
0
1
2


sensor station
start state
goal state
system states
state estimates
position y
position x
3
2
1
0
−1
−2
−1
0
1
2

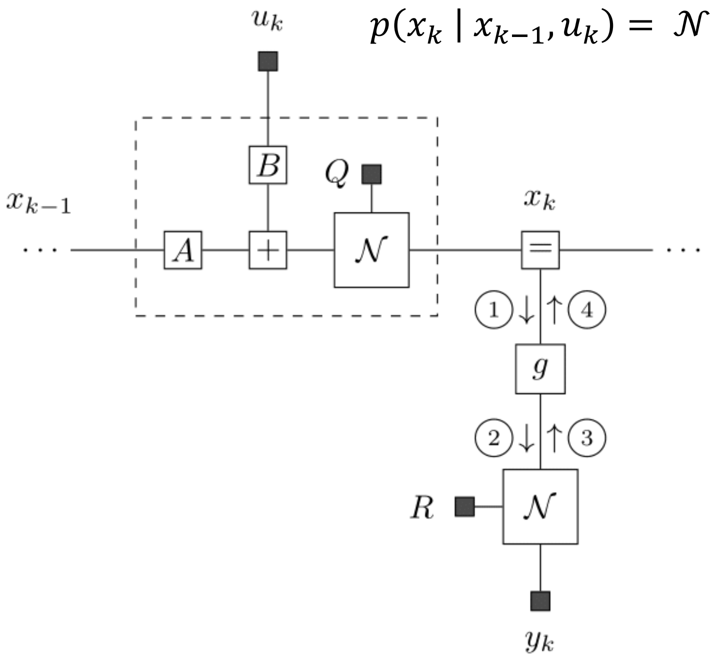
Nonlinear sensing

- Consider a nonlinear observation function: $g(x_k) = \begin{bmatrix} \sqrt{x_1^2 + x_2^2} \\ \arctan(x_2, x_1) \end{bmatrix}$





Model specification



$$p(x_k \mid x_{k-1}, u_k) = \mathcal{N}(x_k \mid Ax_{k-1} + Bu_k, Q)$$

$$1 \downarrow = \mathcal{N}(x_k | \bar{m}_k, \bar{S}_k)$$

Gaussian approximation:

$$p(y_k, x_k) \approx \mathcal{N}\left(\begin{bmatrix} x_k \\ y_k \end{bmatrix} \middle| \begin{bmatrix} \bar{m}_k \\ \mu_k \end{bmatrix}, \begin{bmatrix} \bar{S}_k & \Gamma_k \\ \Gamma_k^\top & \Sigma_k \end{bmatrix}\right)$$

$$2 \downarrow = \mathcal{N}(y_k | \mu_k, \Sigma_k)$$

Inference

- Standard Bayesian filtering for inferring states:

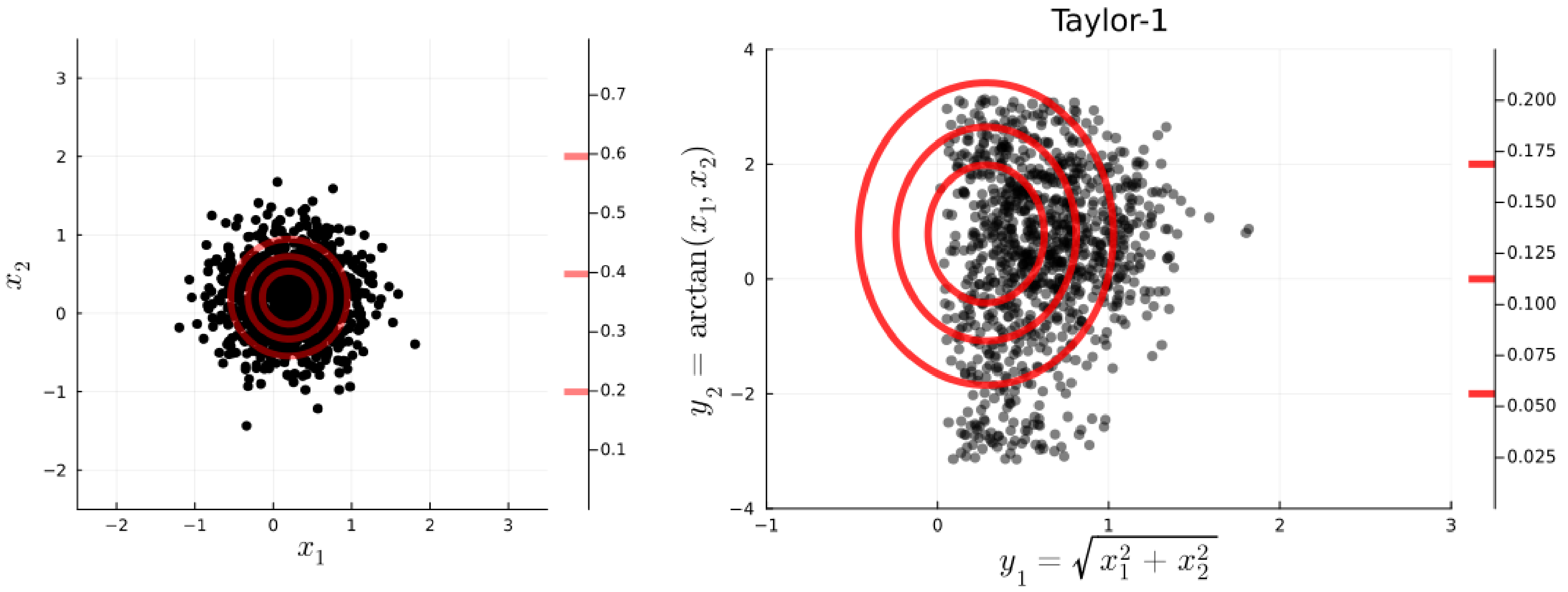
$$p(x_k|\mathcal{D}_k) = \underbrace{\frac{p(y_k|x_k)}{p(y_k|u_k,\mathcal{D}_{k-1})}}_{\uparrow \textcircled{4}\ \text{correction}} \underbrace{\int p(x_k|x_{k-1},u_k)p(x_{k-1}|\mathcal{D}_{k-1})dx_{k-1}}_{\text{prediction}\ \textcircled{1}\downarrow}$$

- Expected free energy for inferring controls at $t > k$:

$$\mathcal{G}(u_t) = \underbrace{\mathbb{E}_{p(y_t|u_t)}\left[\ln\frac{p(y_t|u_t)}{p(y_t|y_*)}\right]}_{\text{risk}} + \underbrace{\mathbb{E}_{p(y_t,x_t|u_t)}\left[-\ln\frac{p(y_t,x_t|u_t)}{p(x_t|u_t)}\right]}_{\text{ambiguity}}$$

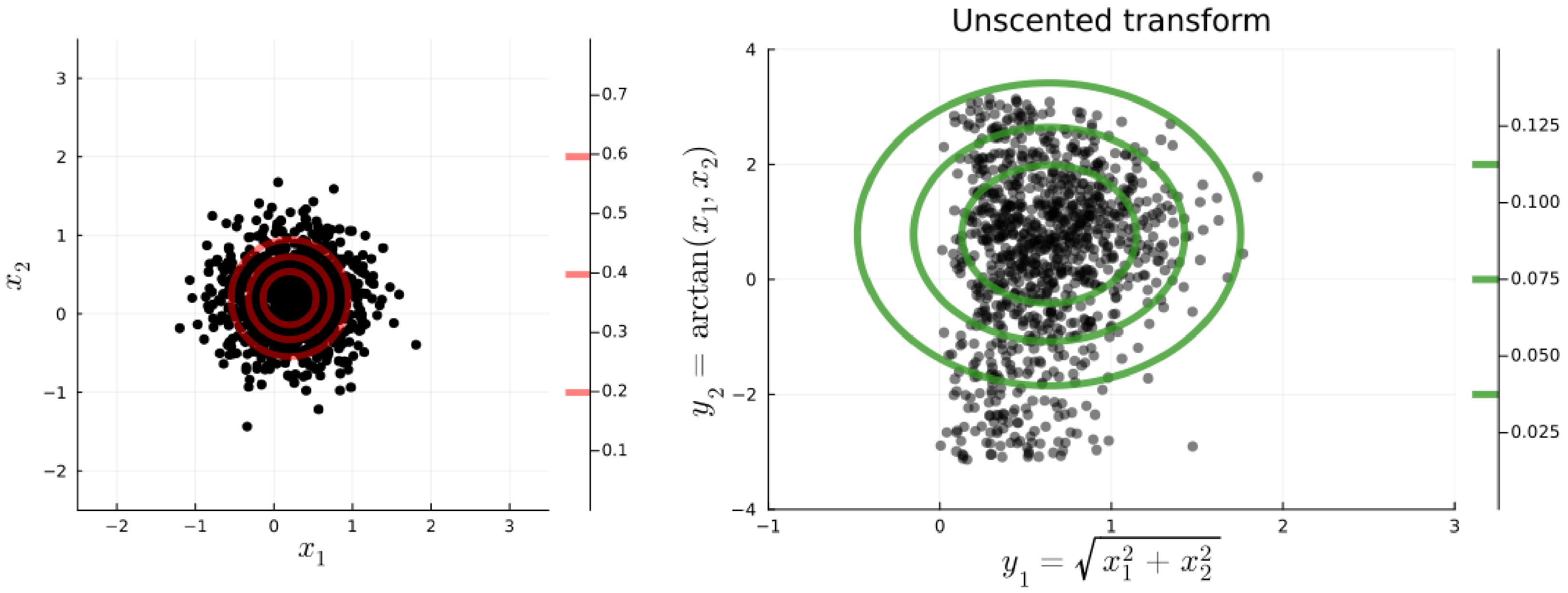
TU/e

Gaussian approximation: 1st – order Taylor



- Theorem: ambiguity $= \frac{1}{2}(D_y \ln 2\pi e + \ln|R|)$

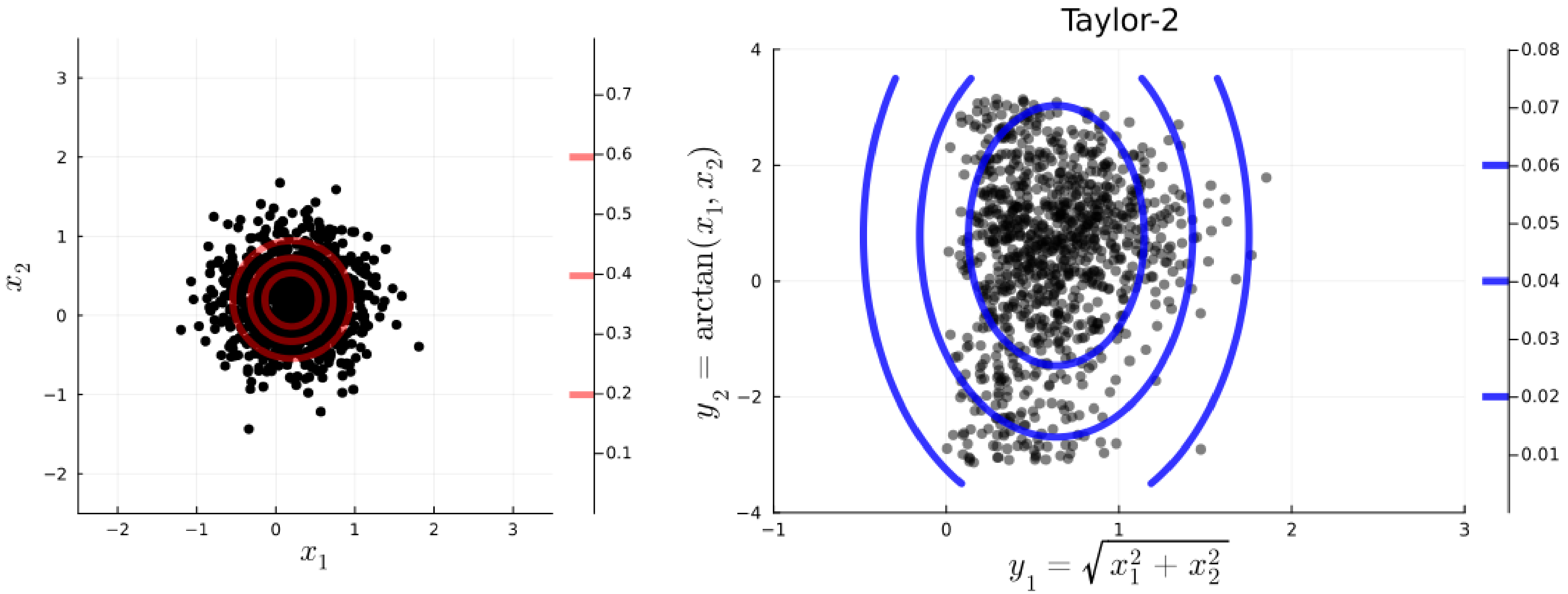
Gaussian approximation: unscented transform



- Theorem: ambiguity $= \frac{1}{2}(D_y \ln 2\pi e + \ln|R|)$

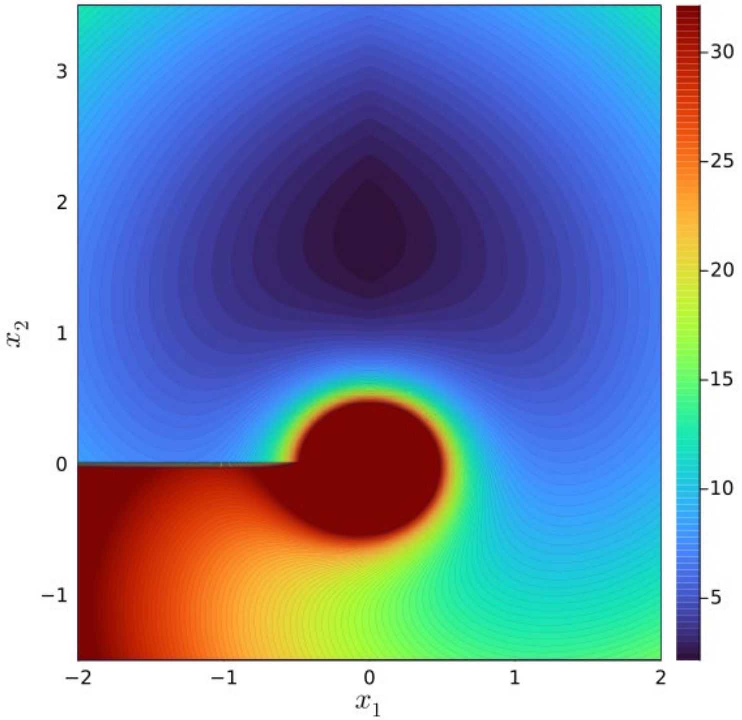
TU/e

Gaussian approximation: 2nd-order Taylor

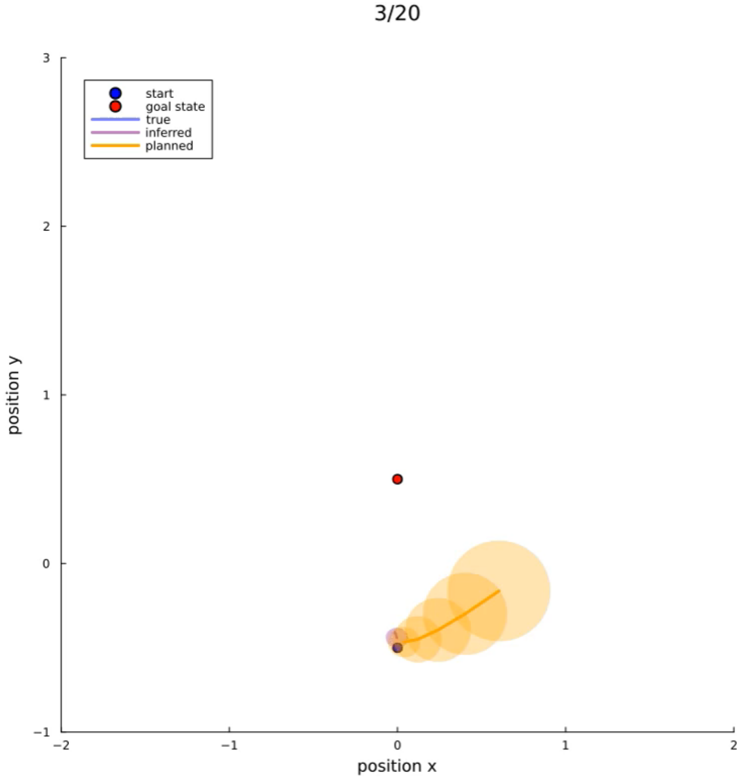


- Theorem: ambiguity $= \frac{1}{2}\left(D_y \ln 2\pi e + \ln\left|R + \frac{1}{2}\sum_{i,j}^{D_y} e_i e_j^\top tr\left(G_{xx}^{(i)}(\bar{m}_k)\bar{S}_k G_{xx}^{(j)}(\bar{m}_k)\bar{S}_k\right)\right|\right)$

Experiment



x_1
x_2



3/20
start
goal state
true
inferred
planned
position x
position y

Discussion

- One the one hand, it is surprising to have a non-zero ambiguity in a fully Gaussian model.
  - Differential entropy is invariant to translation (i.e., does not depend on mean).
- But it is not surprising that a curvature-sensitive approximation affects ambiguity.
  - Differential entropy is not invariant to scaling (i.e., depends on the covariance).

Thank you

Code:

Open PhD position:

Analysis of Bayesian Intelligent Autonomous Systems

Extra slides