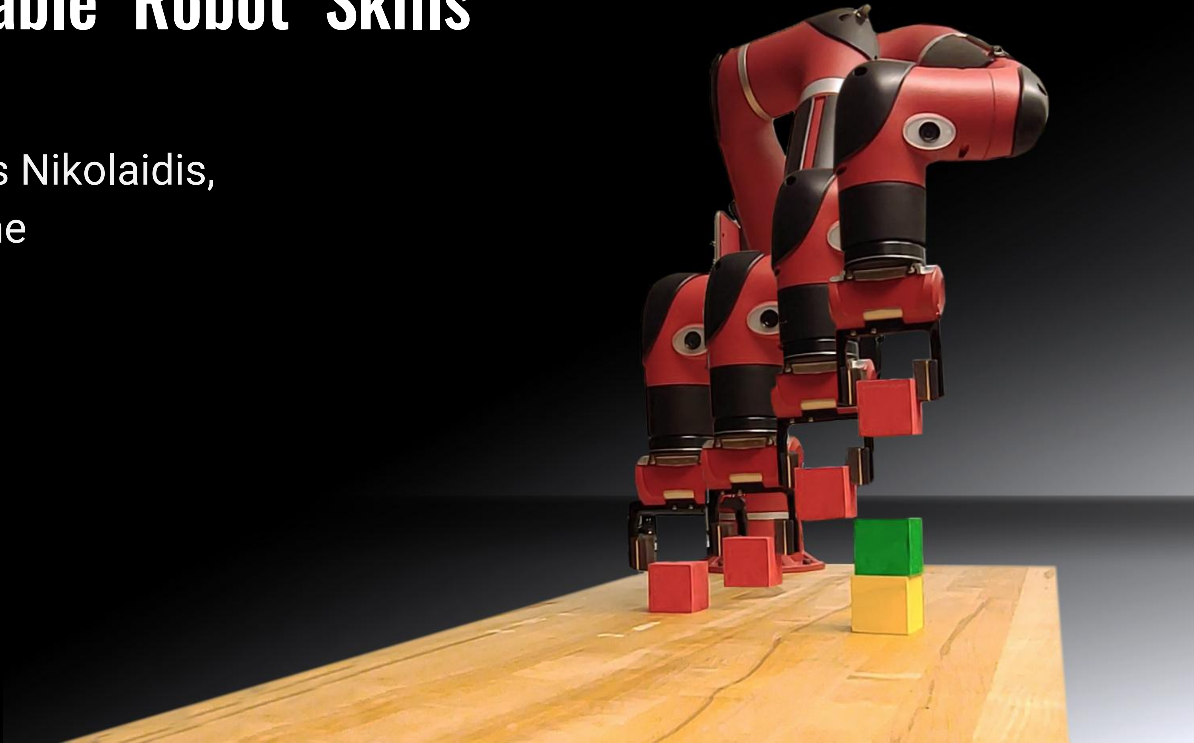
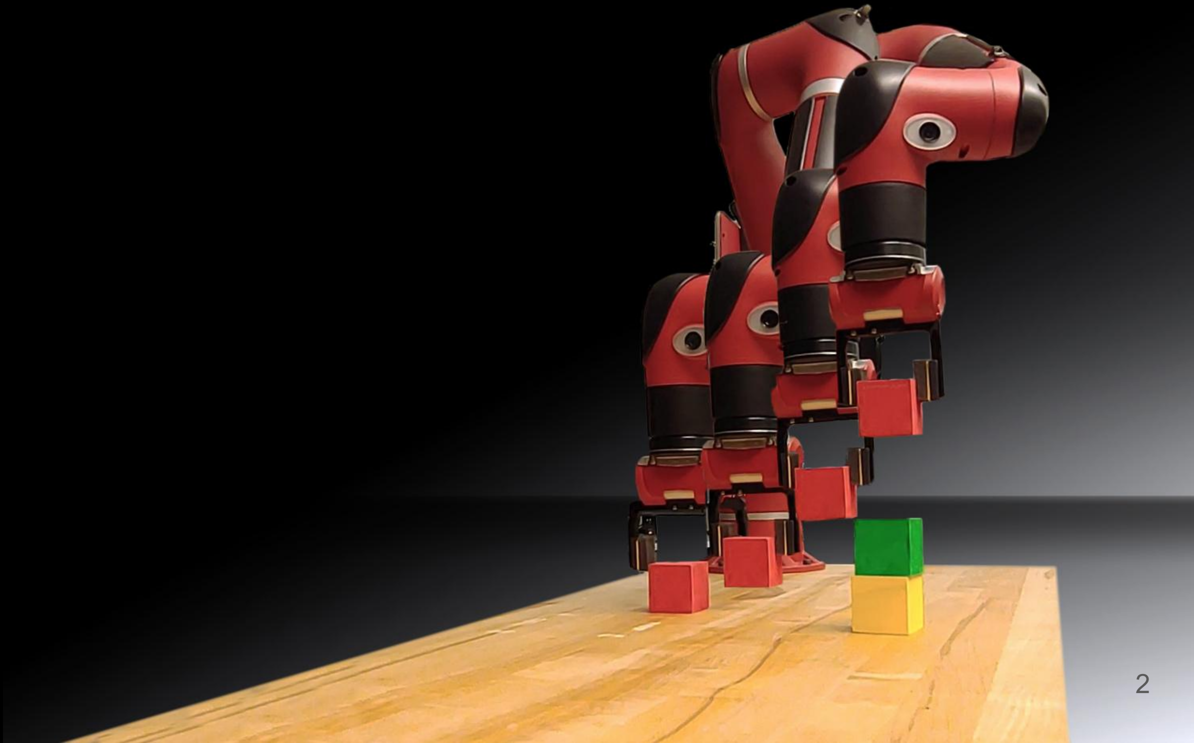
Auto-conditioned Recurrent Mixture Density Networks for Learning Generalizable Robot Skills

Hejia Zhang, Eric Heiden, Stefanos Nikolaidis,
Joseph J. Lim, Gaurav S. Sukhatme

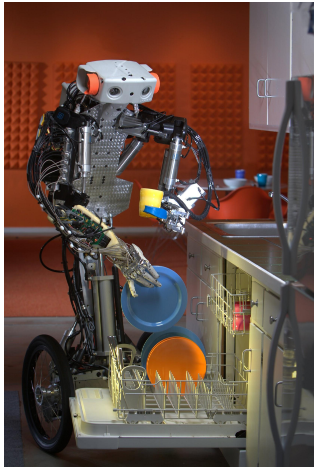


Introduction

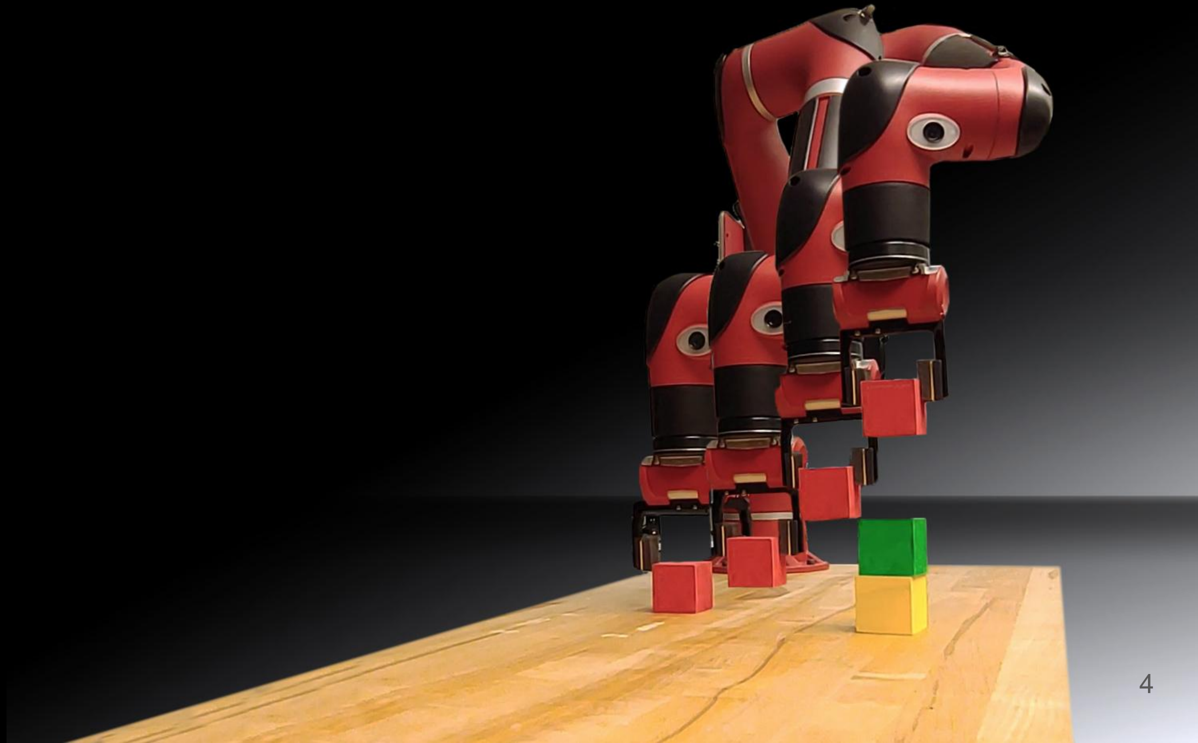


Introduction

- learn generalizable robot skills by imitation learning
- learn state-transition model (STM) to perform tasks with unseen goals
- perform tasks from high-level descriptions
- plan tasks with longer time horizons than the demonstrated tasks
- based on auto-conditioning technique and Recurrent Mixture Density Network (MDN)
- combinable with other methods, e.g. Trajectory Optimization, Inverse Dynamics Models



Architecture



State Transition Model (STM):

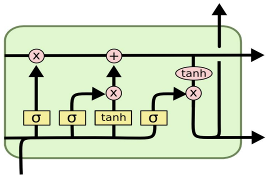
Two requirements for robot skill models:

- Remember long state sequences (history)
- Capture underlying multimodal nature of real world (e.g., different solutions for the same task, human motion prediction)

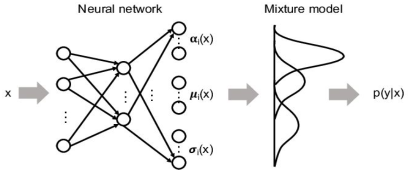
State:
(joint angles, task input, task description)

Recurrent Neural Network

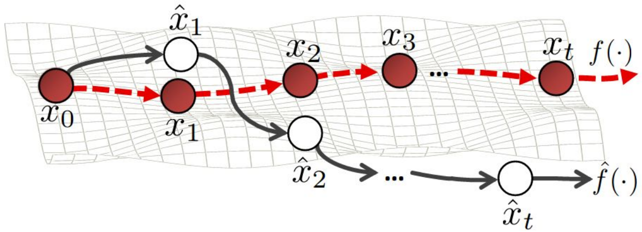
Mixture Density Network



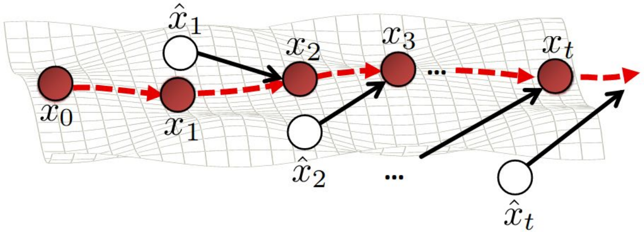
Recurrent Mixture Density Network



Train RNNs via Auto-conditioning



(a) Forward simulation of learned model (gray) introduces error at each prediction step compared to the true time-series (red)



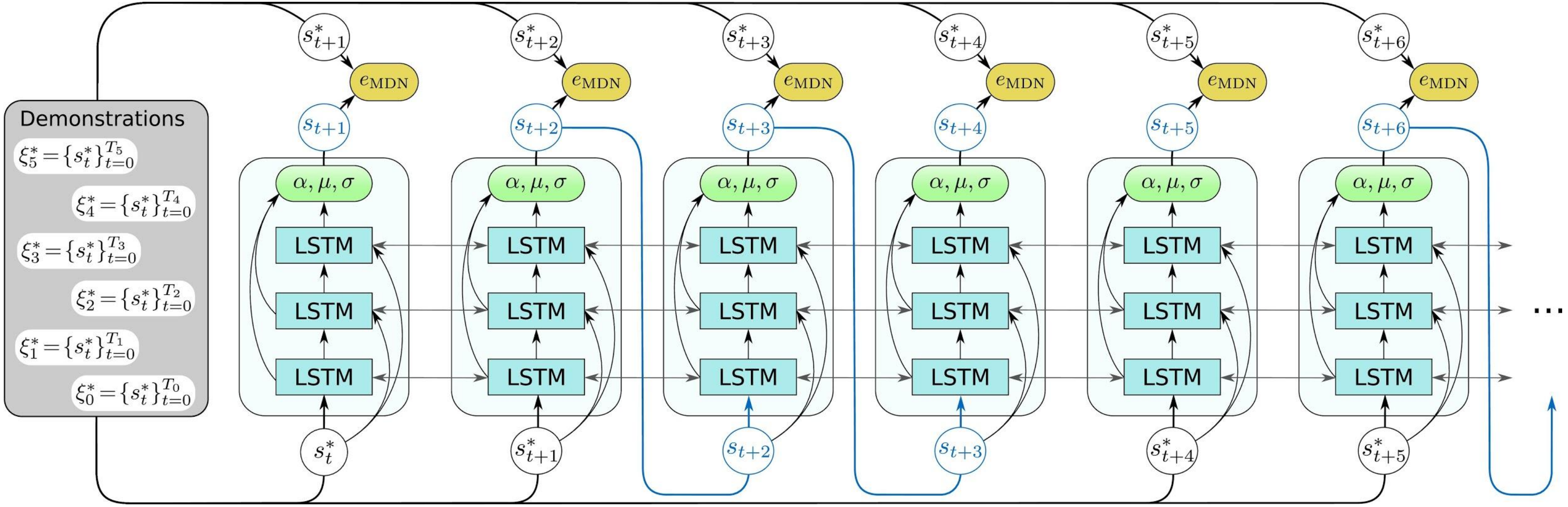
(b) Data provides a demonstration of corrections required to return back to proper prediction

Figure 5: Sample frames from a 300+ second generated sequence. Note that no sequence in the training set exceeds 30 seconds of contiguous motion.

Auto-Conditioned Recurrent Networks for Extended Complex Human Motion Synthesis. Yi Zhou, Zimo Li, Shuangjiu Xiao, Chong He, Zeng Huang, Hao Li. ICLR 2018.

Improving multi-step prediction of learned time series models. Arun Venkatraman, Martial Hebert, J. Andrew Bagnell. AAAI 2015.

Architecture



Trajectory Optimization

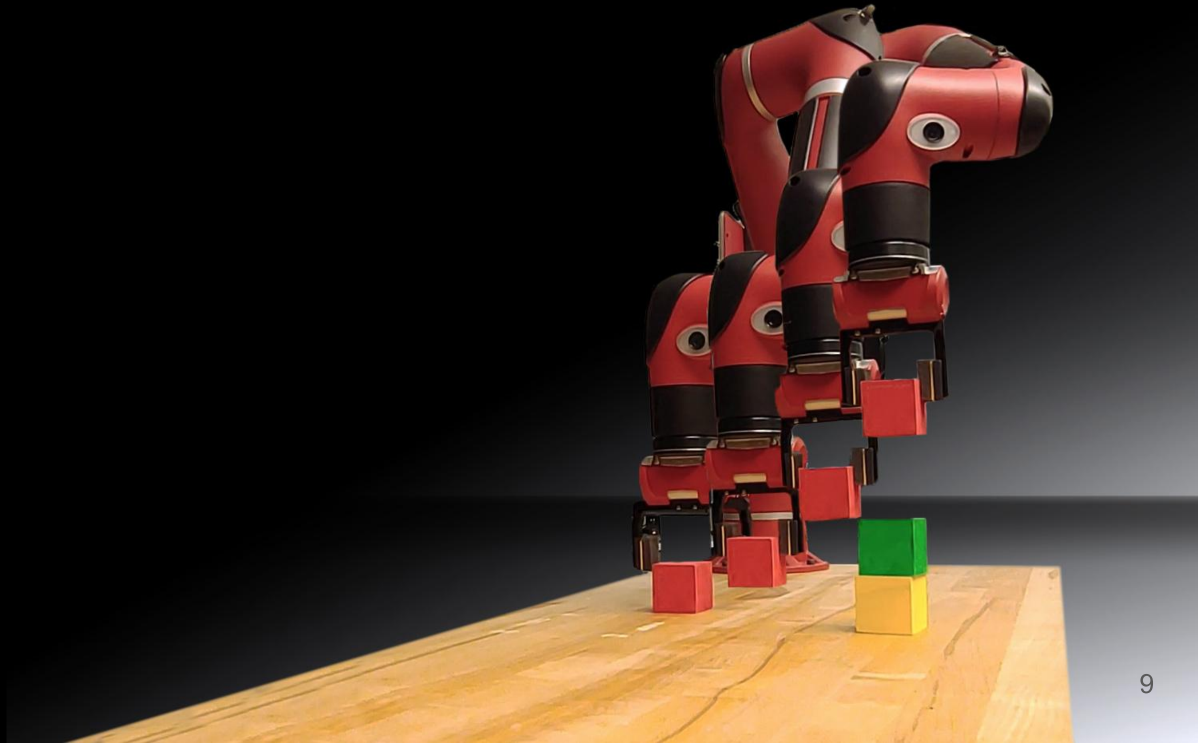
Smooth trajectory by minimizing the objective

$$\{\mathbf{q}_t^*\}_{t=1}^T = \underset{\{\mathbf{q}_t\}_{t=1}^T}{\arg\min} V(\{\mathbf{q}_t\}_{t=1}^T)$$

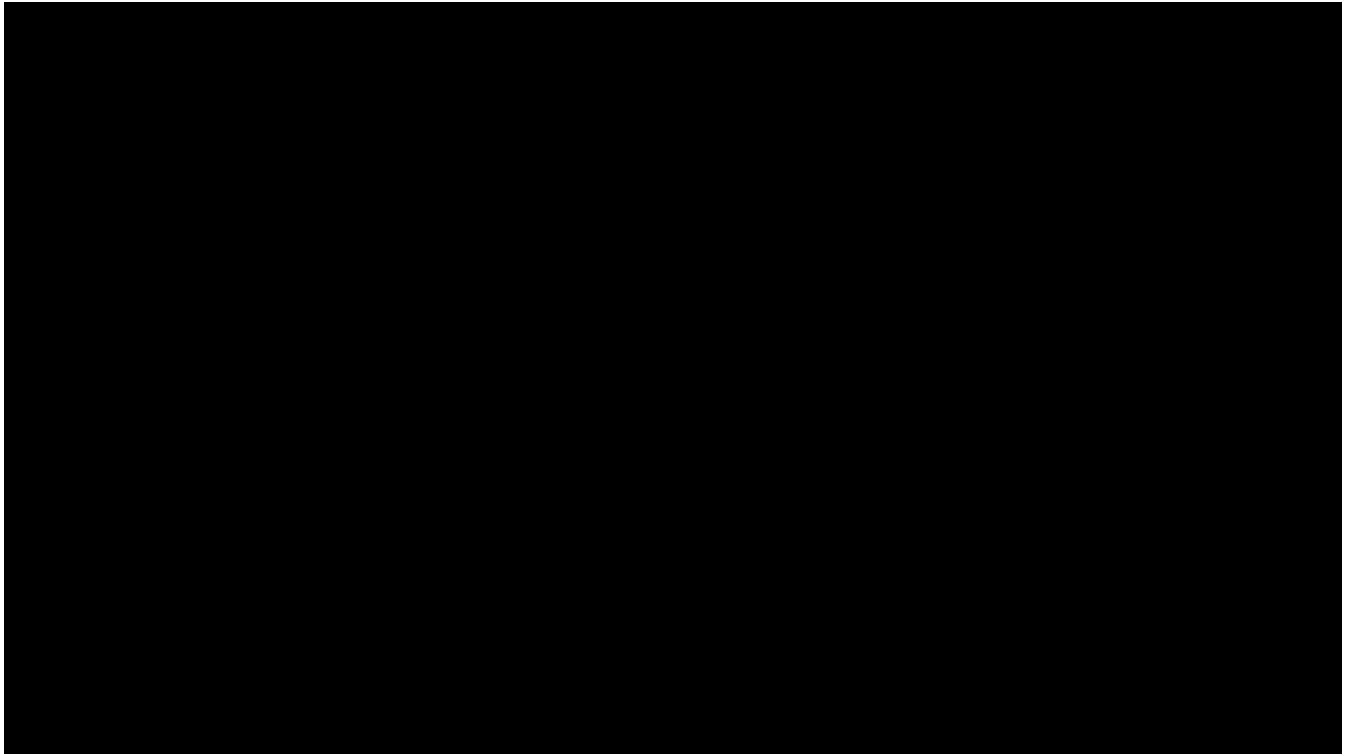
where

$$V(\{\mathbf{q}_t\}_{t=1}^T) = \sum_{t=1}^{T-1} \|\mathbf{q}_t - \tilde{\mathbf{q}}_t\|_2^2 + \gamma \|\mathbf{q}_{t+1} - \mathbf{q}_t\|_2^2$$

Experiments



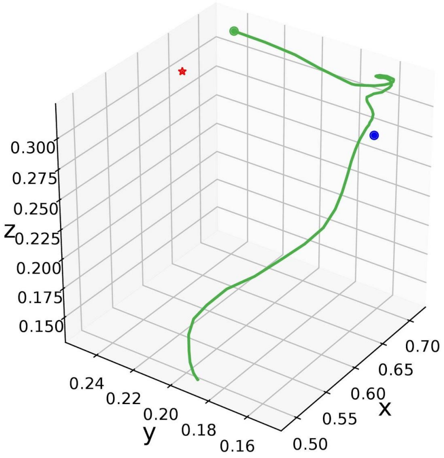
Experiment - Stacking blocks



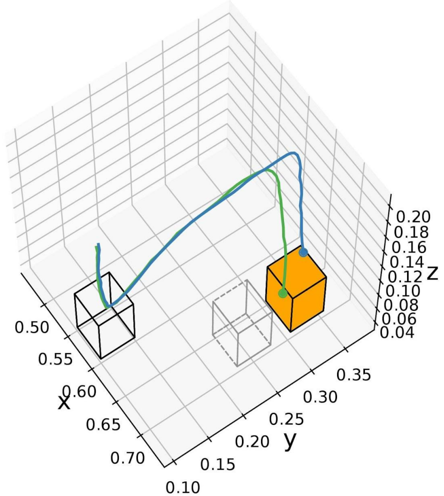
Experiment - Drawing circles

Experiment - Adaptability

Reaching



Pick & Place

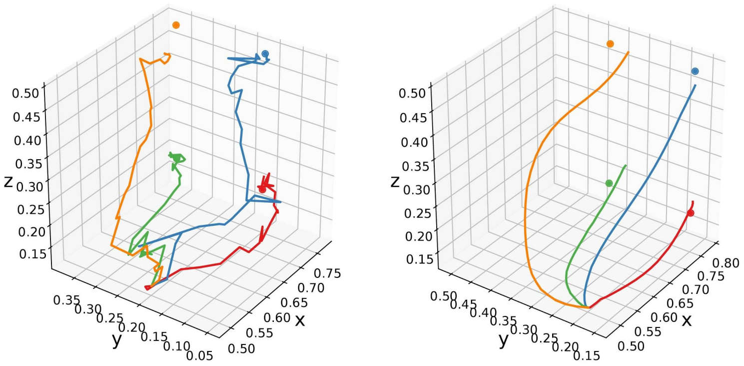


The goal is changed at the middle of each task execution. The plot shows how our model can adapt to changing goals and still works beyond the planning horizon of its demonstrations.

Experiment - Combine with other methods

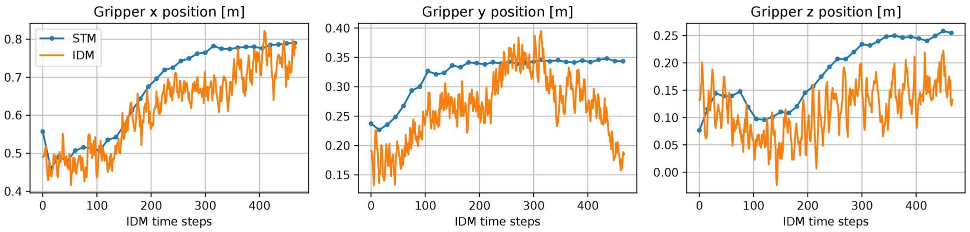
- trajectory optimizer for smoothness and precision (goal-based)
- inverse dynamics model (IDM) for efficient sim-2-real transfer

Reaching to 4 goals



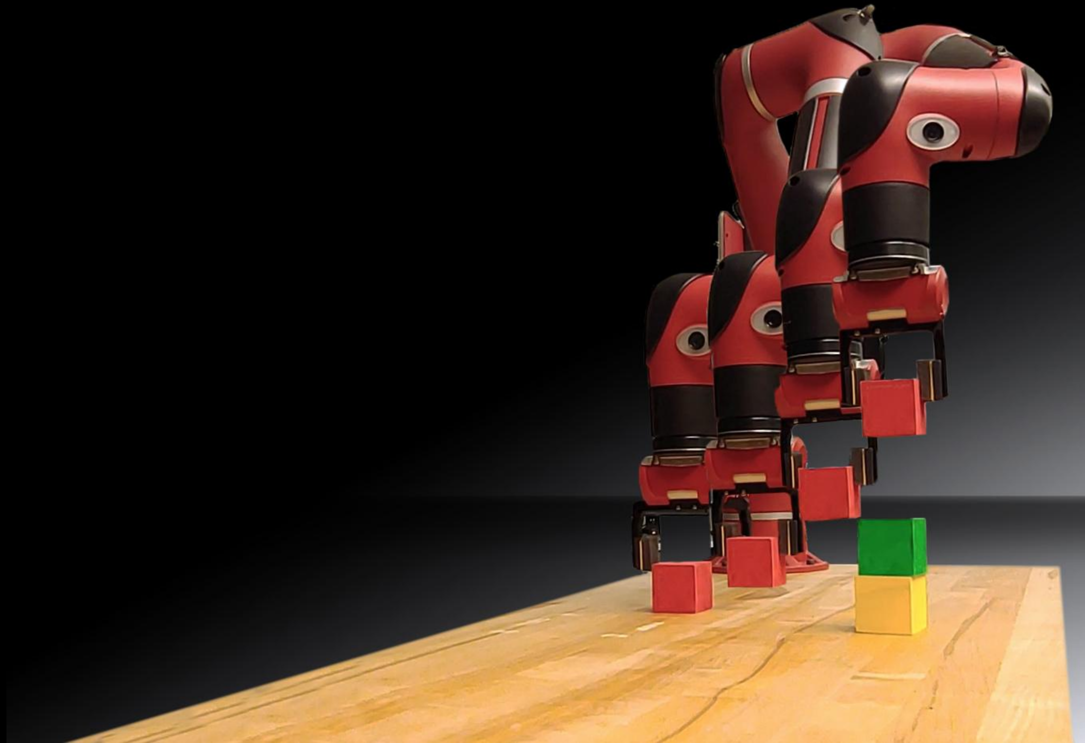
Trajectories before and after smoothing

Reaching to 1 goal



Combination with inverse dynamics model

Conclusion



Conclusion

State:
(joint angles, human motions, task input, task description)

Deeper insight into our neural network structure:

- **Assumption 1**: Every single task can be solved in several ways.
- **Assumption 2**: Different phases of a single task governed by different mixture Gaussian components
  (e.g. approaching, grasping, placing for pick-and-place tasks)
    "How do Mixture Density RNNs Predict the Future". Kai Olav Ellefsen, Charles Patrick Martin, Jim Torresen. Arxiv preprint, 2019.

Future directions:

- Investigate roles of individual Gaussians of MDN applied to learning robot skills
  (based on Ellefsen's work)
- Generalize towards more complex tasks with human teammates
- Connect with trajectory optimization methods
  (optimize over variety of dynamic and task-based criteria)

Auto-conditioned Recurrent Mixture Density Networks for Learning Generalizable Robot Skills

Hejia Zhang, Eric Heiden, Stefanos Nikolaidis,
Joseph J. Lim, Gaurav S. Sukhatme