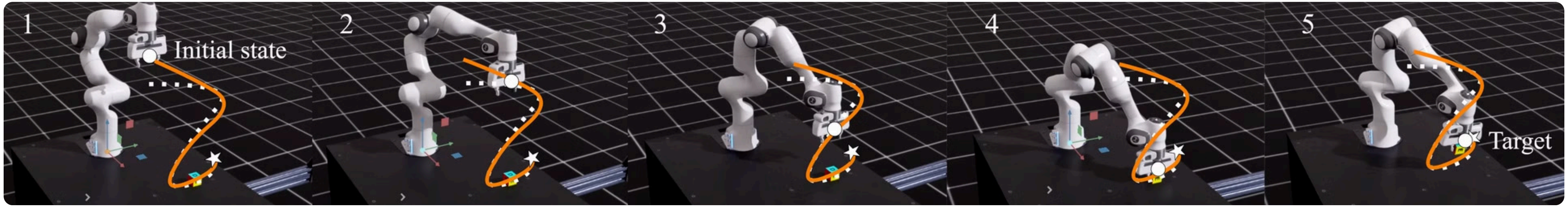
Contractive Dynamical Imitation Policies for Efficient Out-of-Sample Recovery

Paper presentation

ICLR 2025, Singapore

Amin Abyaneh, Mahrokh G. Boroujeni, Hsiu-Chin Lin, and Giancarlo Ferrari-Trecate



ICLR

McGill

EPFL

Imitation Learning

Problem Setup

Expert data

Collection of optimal expert trajectories

Supervised learning

Planning policy

Policy to plan the next state, or current action

Objective

- Learn policies to mimic the expert behavior
- Recovery from unseen states
- Formal guarantees for recovery

Imitation via Dynamical System

Expert trajectories as solutions to an

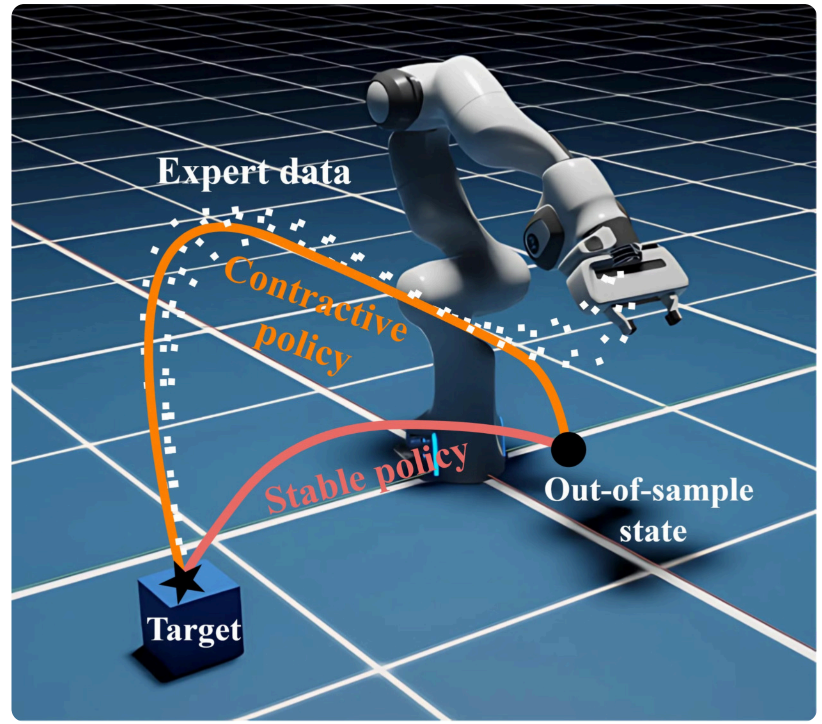
ordinary differential equation (ODE)

Stability → solutions converge to an equilibrium

Contractivity → all solutions contract

- Stability is a by-product

Contraction improves the transient behavior



Stories from the Literature

CSDP[1]

- [ ] Limited representation power
- [ ] Complex optimization
- [x] ~~SOS scaling issues~~

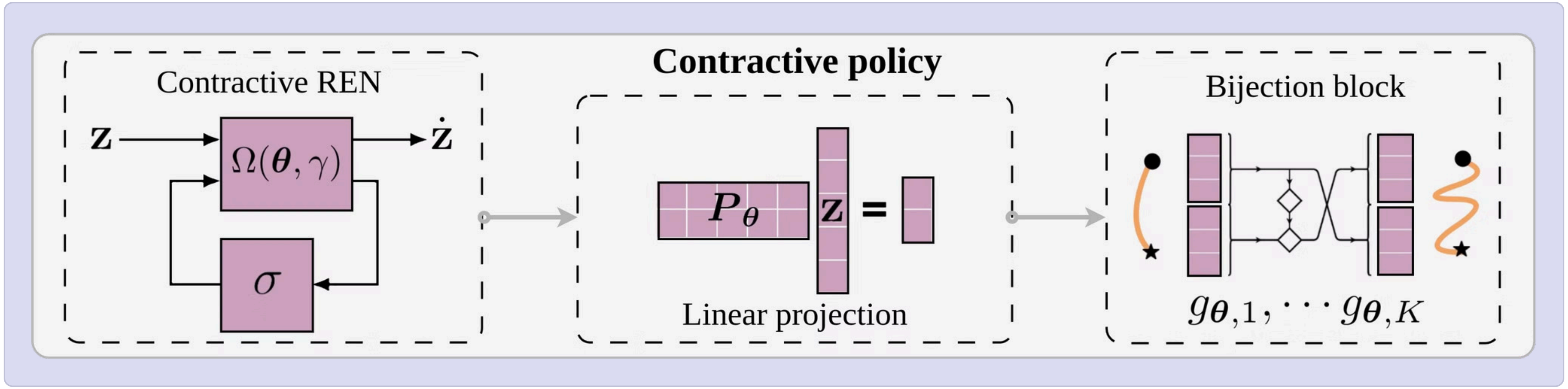
SCDS (ours)

- [ ] Variable contraction rate
- [ ] Higher dimensional tasks
- [ ] State-only formulation

NCDS[2]

- [ ] Second order ODE solutions
- [x] ~~Lower-dimensional latent space~~

Expressive Contractive Policies



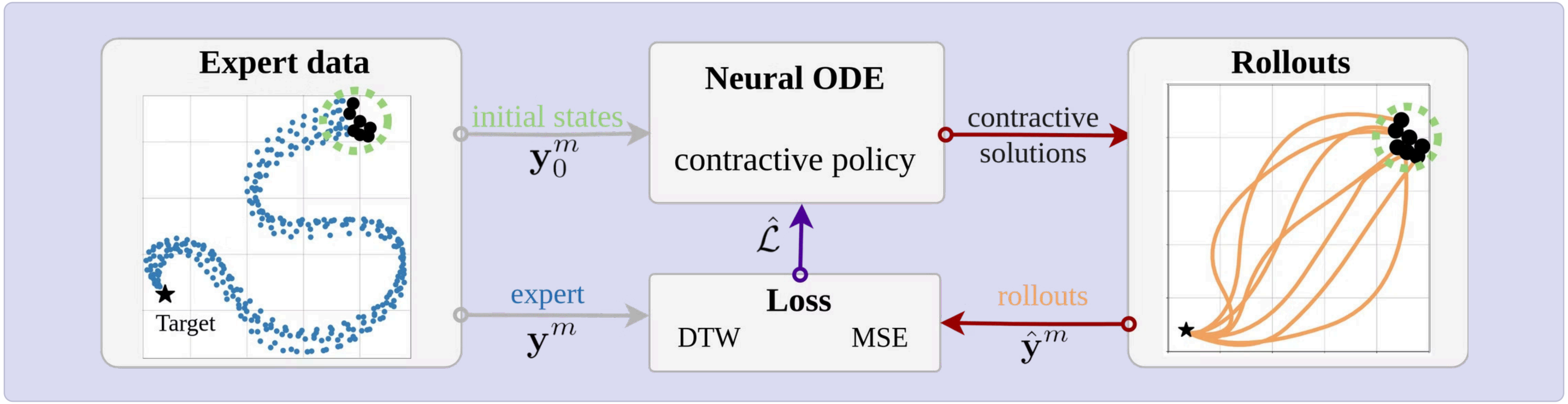
REN (latent) dynamics

- Contractive for **any** choice of parameters
- Learnable (and adjustable) contraction rate

Output transformation

- Trainable linear projection
- Invertible contraction-preserving coupling layers

Training with Neural ODEs



Differentiable ODE Solutions

Initial value problem to generate differentiable trajectories.

Trajectory Space Loss

Policy rollouts are compared with expert data using dynamic time warping or MSE.

Optimization Problem

Optimal parameters are learned by minimizing the empirical loss.

Out-of-Sample Error Guarantees

1

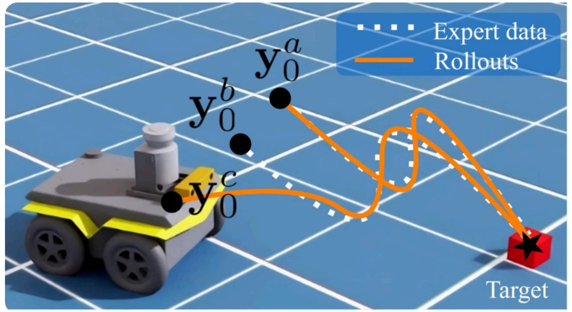
Bounded initial states

- Initial state lies within a multi-focal ellipse region
- Focal points at the initial conditions in the dataset

2

Upper-bounding the loss

- Weighted sum of MSE and
- Uncertainty in the initial state



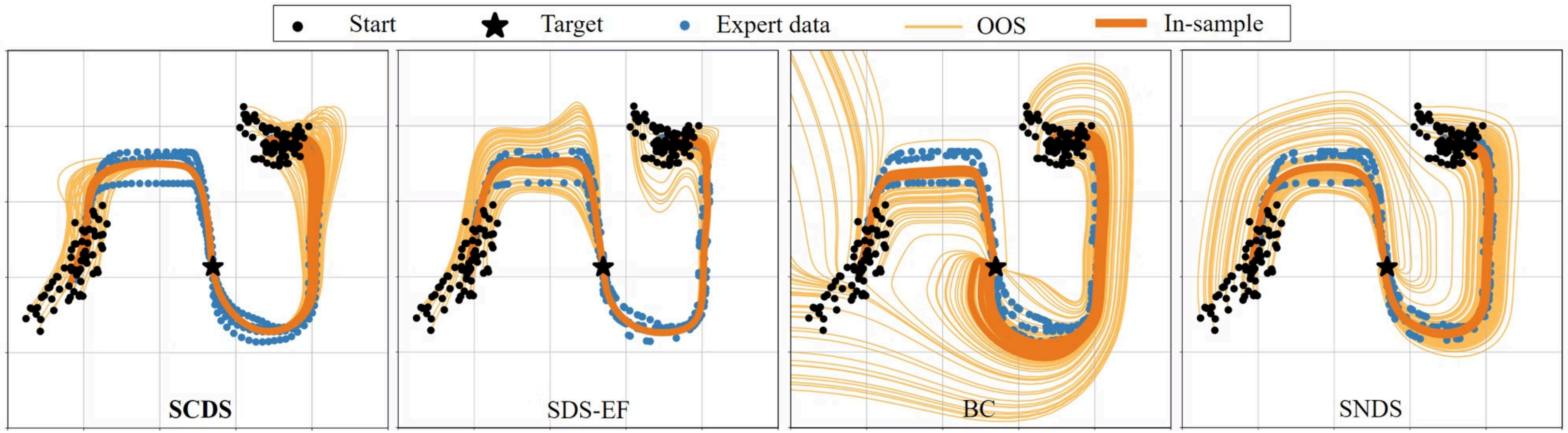
Theorem.

Loss for a new trajectory ≤ Loss of training trajectories (weighted average) + Uncertainty factor

Experiments

LASA dataset

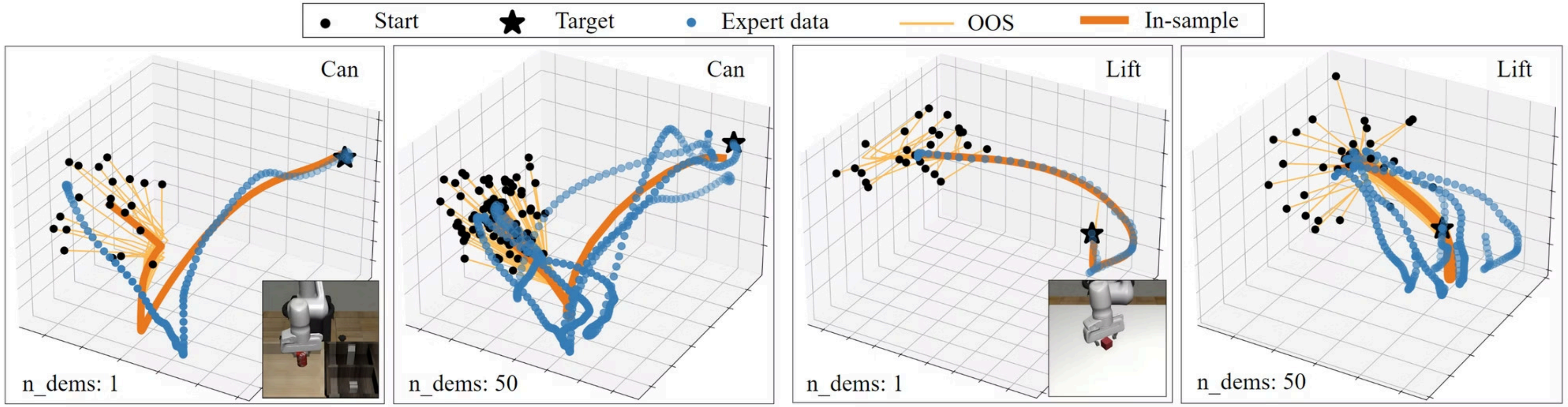
- [ ] OOS rollouts reliably contract towards expert demonstrations with high precision.



Experiments

Robomimic dataset

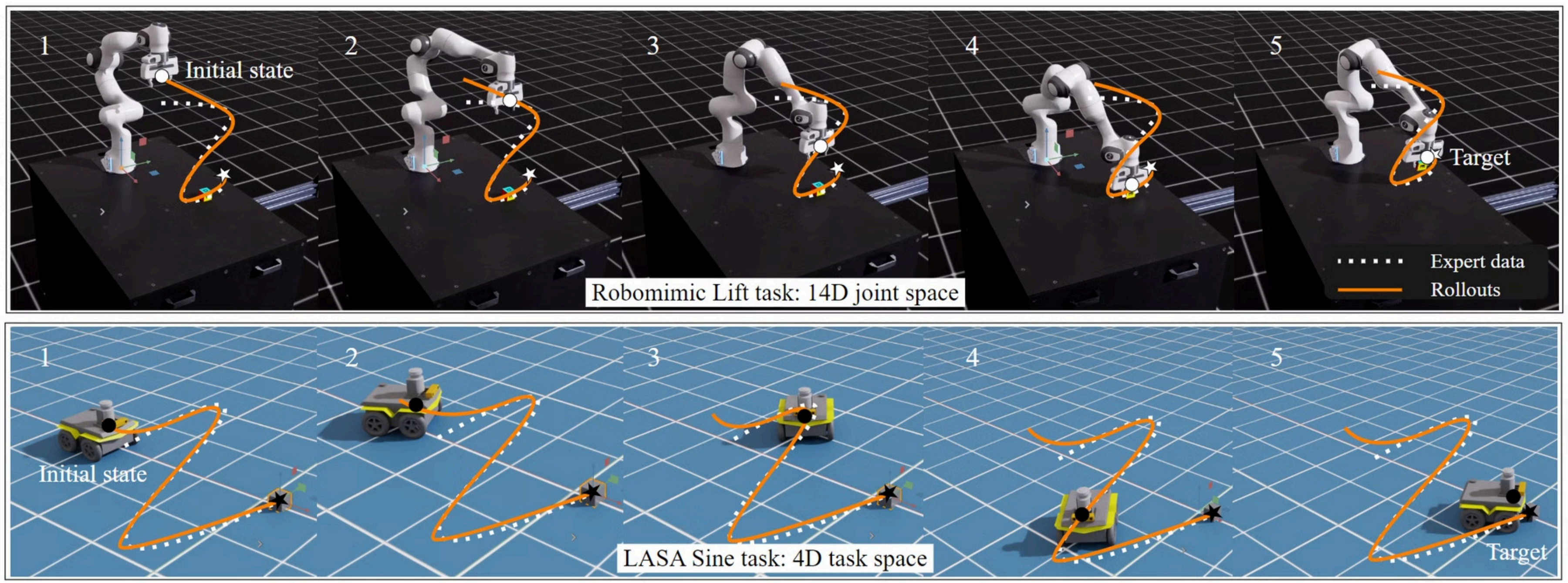
- [ ] Learning in higher dimensional state spaces can be achieved efficiently.



Experiments

Simulation deployment

- Planning for manipulator and wheeled robots in simulation.



Experiments

Simulation results

- Substantial improvement in out-of-sample initial states compared to baselines.
- Out-of-sample recovery error remains within the calculated bounds!

Expert	LASA-2D		LASA-4D		Robomimic-6D		Robomimic-14D	
Metric	MSE	soft-DTW	MSE	soft-DTW	MSE	soft-DTW	MSE	soft-DTW
SNDS	0.02 ± 0.01	0.72 ± 0.14	0.03 ± 0.01	1.04 ± 0.19	0.65 ± 0.55	1.26 ± 0.70	2.42 ± 1.37	4.15 ± 0.92
BC	0.04 ± 0.02	0.98 ± 0.12	0.05 ± 0.03	1.48 ± 0.16	0.56 ± 0.32	1.88 ± 0.20	1.75 ± 0.22	4.86 ± 0.58
SDS-EF	0.03 ± 0.01	0.85 ± 0.13	0.05 ± 0.02	1.10 ± 0.15	0.50 ± 0.28	1.01 ± 0.66	3.30 ± 0.75	5.65 ± 0.58
SCDS	0.02 ± 0.01	0.65 ± 0.05	0.03 ± 0.01	0.72 ± 0.12	0.56 ± 0.22	1.05 ± 0.37	1.68 ± 0.45	4.10 ± 0.40
SNDS	2.73 ± 1.67	6.91 ± 1.46	3.65 ± 2.12	9.85 ± 0.63	1.58 ± 0.93	3.27 ± 1.88	6.88 ± 2.16	12.17 ± 2.70
BC	8.63 ± 4.05	16.25 ± 5.27	19.25 ± 7.34	27.48 ± 6.83	11.05 ± 7.41	19.94 ± 9.57	44.98 ± 15.11	37.82 ± 14.13
SDS-EF	1.78 ± 0.34	7.13 ± 1.51	2.33 ± 0.47	10.21 ± 1.98	1.15 ± 0.81	2.67 ± 1.40	11.10 ± 2.10	10.44 ± 1.66
SCDS	0.32 ± 0.15	1.72 ± 0.54	1.09 ± 0.21	2.58 ± 0.30	0.89 ± 0.26	1.71 ± 0.53	2.68 ± 0.65	6.27 ± 0.48
$\mathcal{L}_{\text{ub}}^{\text{MSE}}$	0.49 ± 0.07		1.33 ± 0.14		1.43 ± 0.22		2.91 ± 0.45	

Made with Gamma

Conclusion

Summary

- [x] ~~Efficient OOS recovery~~
- [x] ~~Variable contraction rate~~
- [x] ~~Contractive-by-design~~

Future Work

- [x] ~~Long-horizon contractive planning~~
- [x] ~~Stochastic contractive policies~~

Thank you!

Find us at the poster session to talk in person 🙂

Correspondence: amin.abyaneh@mail.mcgill.ca