Dynamics Are Learned, Not Told

Semi-Supervised Discovery of Latent Dynamics Geometries For Zero-Shot Policy Adaptation

Zhiming Xu[1] Weitao Zhou[2 3 *] Xianghui Pan[1] Nanshan Deng[2 4] Chengju Liu[1] Qijun Chen[1] Chenpeng Yao[1 *]

[1] Tongji University [2] Tsinghua University [3] Simple AI [4] Rimbot

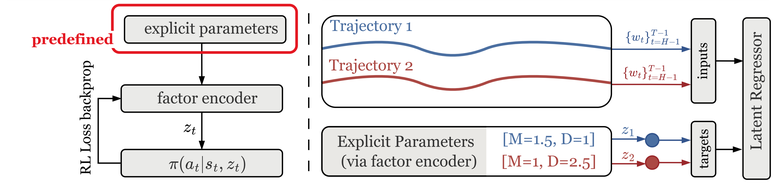
Tongji University · Tsinghua University · Simple AI · Rimbot

ICML International Conference On Machine Learning

Paper · GitHub

Motivation: Beyond Parameter-Centric Dynamics Adaptation

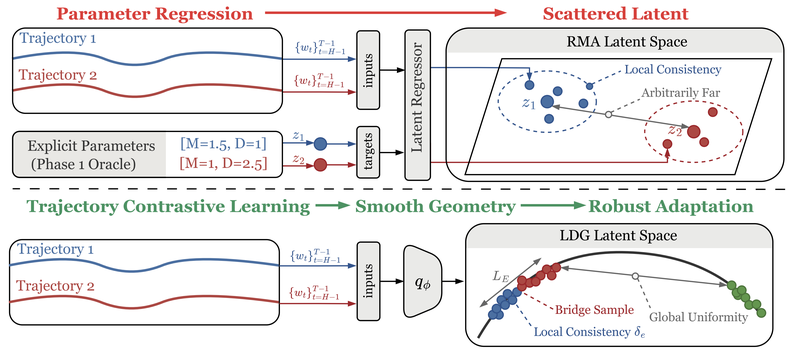
Parameter-Centric (Traditional, e.g., Rapid Motor Adaptation)



Problem I: real robots are subject to unmodeled, coupled and time-varying factors, for which no fixed set of physcial parameters provides a complete description.
Problem II: distinct physical phenomena often induce indistinguishable effects on motion (e.g., increased payload and higher friction all manifest as "resistance to motion").

Can we learn a latent space without relying on privileged dynamics parameters?
——Not only can we, but we should!

The Big Picture: An Outcome-Centric Perspective to Shape Latent Geometries



Theoretical Guarantee

Theorem (Target Domain Regret Bound)

Let $\mathcal{M}_S$ and $\mathcal{M}_T$ be two MDPs differing solely in their transition dynamics. The upper bound for target domain regret is monotonically increasing with respect to trajectory encoder consistency δ_e and smoothness L_E.

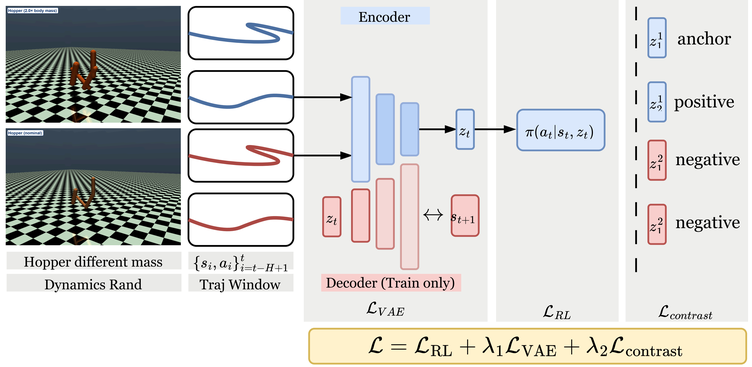
Theorem (Contrastive Minimization of L_E)

If the source and target distributions have sufficient overlap ($\mathbb{P}(S) > \alpha$), then the alignment loss in InfoNCE decomposition strictly bounds L_E by $\sqrt{\mathcal{L}_{align}}$.

Theorem (Gradient Orthogonality to Noise)

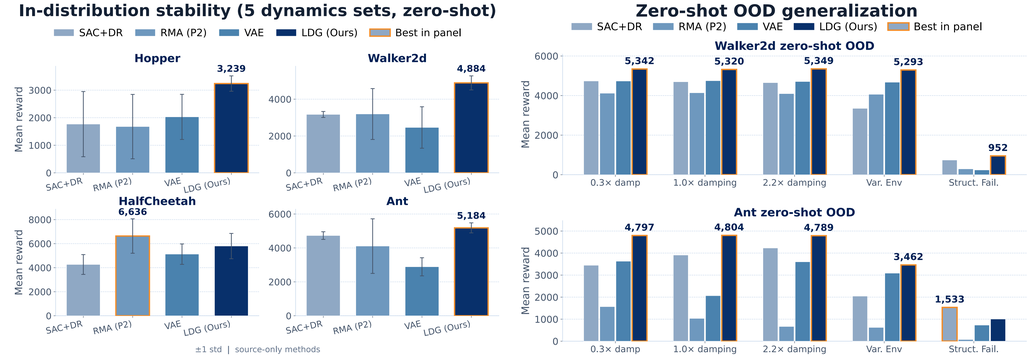
If a trajectory factors into pure dynamics μ and nuisance features s such that $\tau \approx (\mu, s)$, minimizing InfoNCE penalizes sensitivity to noise ($\|\partial E/\partial s\|_F$), forcing an orthogonal, purely dynamics-driven latent space.

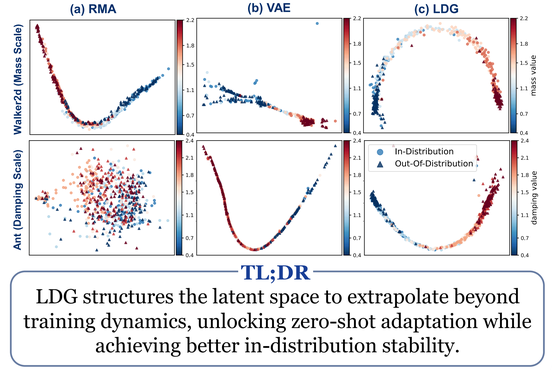
Method Overview: RL + VAE + InfoNCE



$$\mathcal{L} = \mathcal{L}_{\text{RL}} + \lambda_1 \mathcal{L}_{\text{VAE}} + \lambda_2 \mathcal{L}_{\text{contrast}}$$

Empirical Results: In-Distribution Stability and Zero-Shot Adaptation to Out-of-Distribution Domains

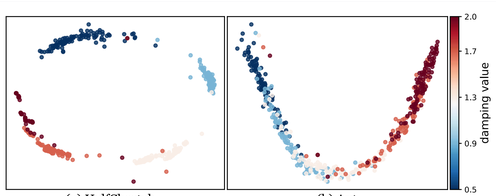




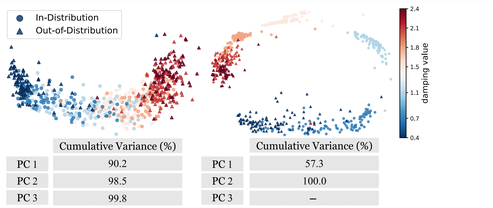
TL;DR

LDG structures the latent space to extrapolate beyond training dynamics, unlocking zero-shot adaptation while achieving better in-distribution stability.

Latent Structure Discovery

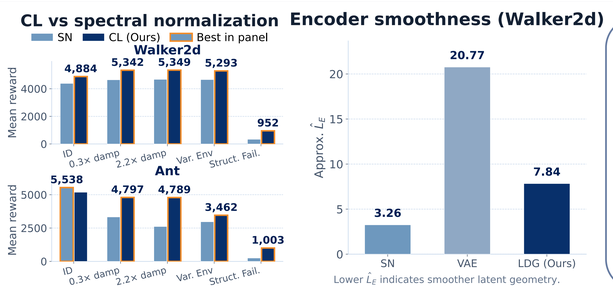


LDG learns functional equivalence: after training on varying mass, the encoder organizes unseen damping variations into a coherent manifold



Physical Isomorphism: the learned latent space exhibits a topological complexity that mirrors the physical constraints of the agent

Apart From Smoothness: Geometry Matters



TL;DR

1. controlling network smoothness is a fundamental principle shared across generative vision and policy adaptation.
2. **naively enforcing smoothness is insufficient:** successful adaptation requires simultaneous enforcement of smoothness and topological structure within the latent space.

Key Takeaways

- **Interaction Over Parameters:** Shape the dynamics latent space entirely through interaction outcomes rather than relying on privileged dynamics parameters.
- **Geometry Drives Generalization:** While smoothness ensures in-distribution stability, it is the underlying geometry that unlocks out-of-distribution generalization.
- **Contrastive Topology:** Contrastive learning bounds the trajectory encoder's Lipschitz constant, yielding a smooth and structured latent space.

Semi-supervised learning on unlabeled trajectories to discover latent dynamics geometries for robust zero-shot policy adapation

Contact: 2251804@tongji.edu.cn, zhouwt801@gmail.com