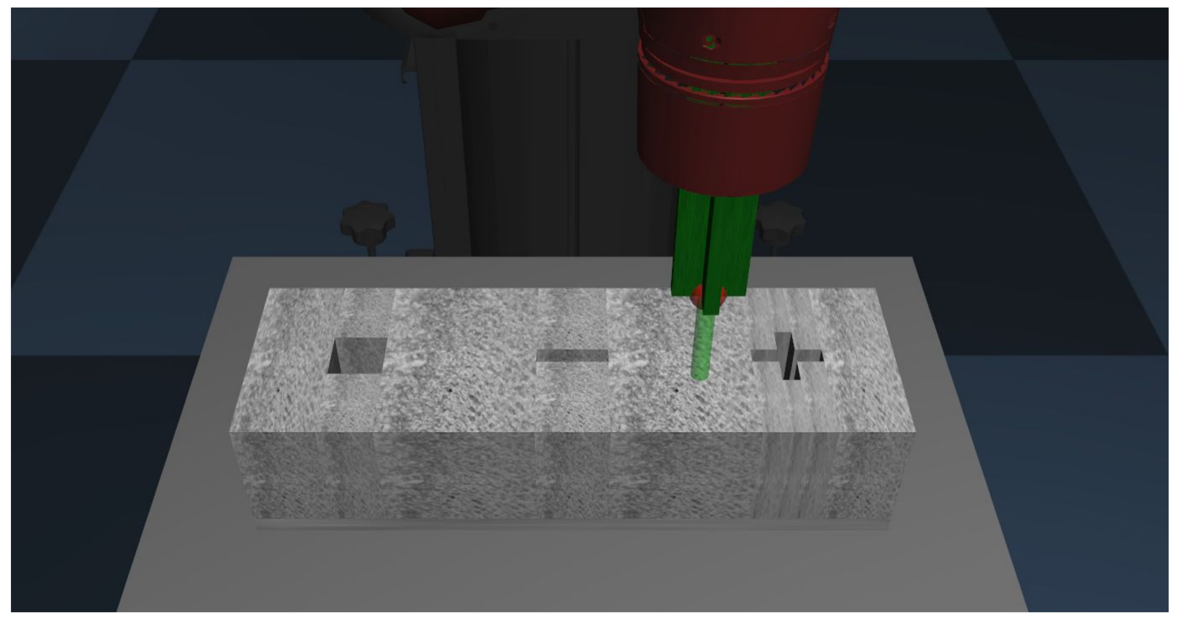
Learning Latent Space Dynamics Models for Planning in Robotics

Peter Zachares



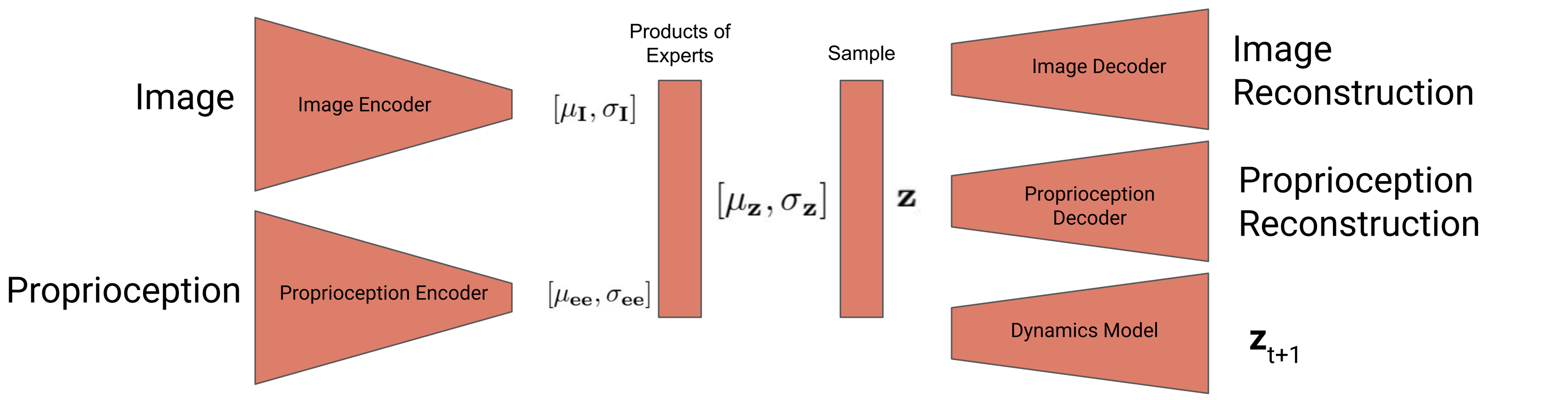
Peg Insertion: Mujoco Simulation Environment

Data collection Method:

- Collected using a heuristic waypoint following policy which accomplished the task of peg insertion
- Each data point (of ~600,000) includes an image, force-torque reading, 6DOF end-effector position and velocity and the action taken at that step in the trajectory

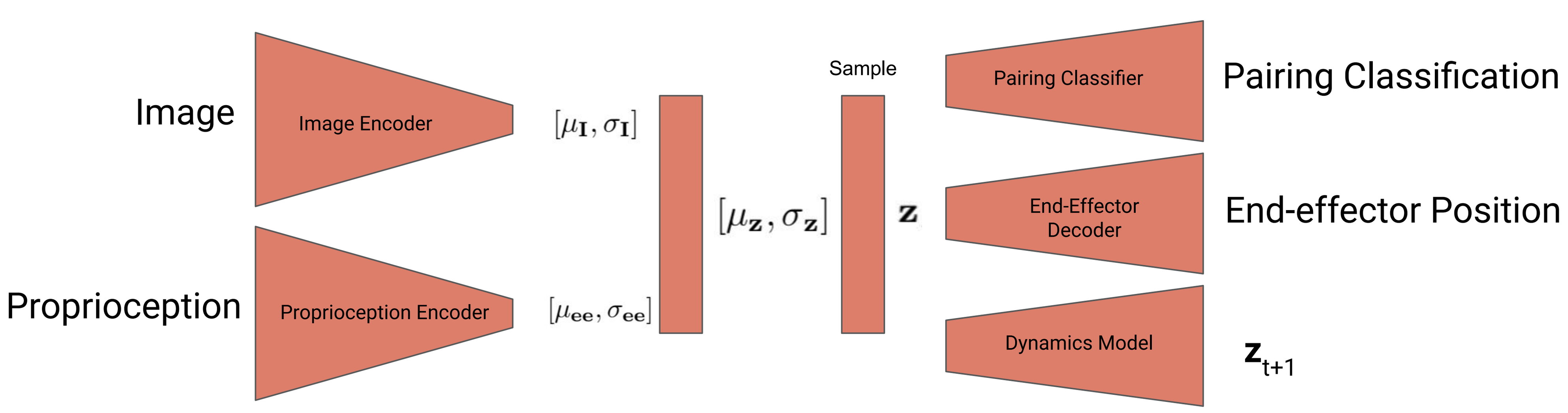
Joint Representation and Dynamics model Learning Method

Model 1: Variational Autoencoder with 6 layer MLP Dynamics Model



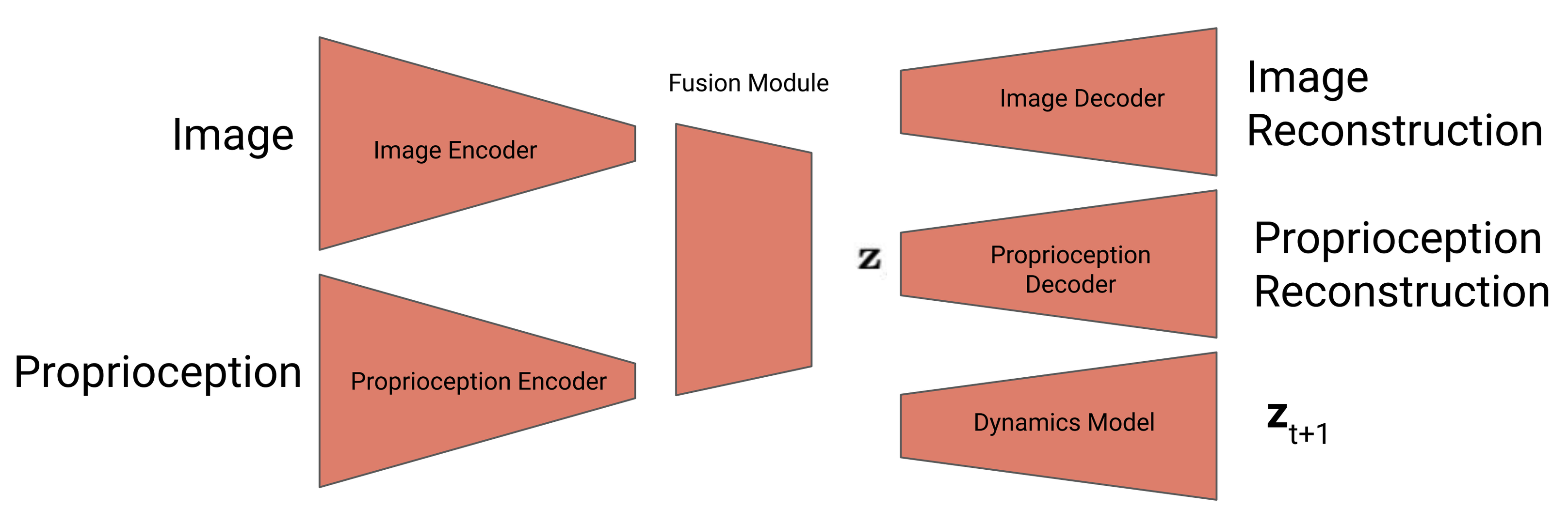
$$Loss = \log(p(image_{pred}|z_t)) + \log(p(prop_{pred}|z_t)) - D_{KL}(q(z_{t+1}|image_{t+1}, prop_{t+1})|p(z_{t+1}|z_t))$$

Model 2: Self-Supervised Variational Encoder with 6 layer MLP Dynamics Model

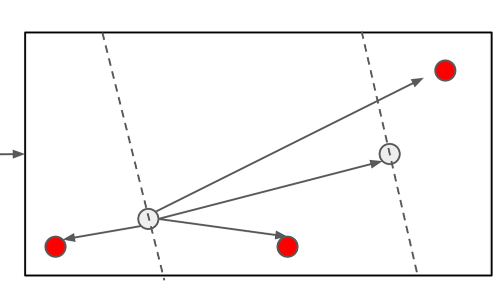


$$Loss = \log(p(pair_{class}|z_t)) + \log(p(eepos_{pred}|z_t)) - D_{KL}(q(z_{t+1}|image_{t+1}, prop_{t+1})|p(z_{t+1}|z_t))$$

Model 3: Relational Autoencoder with 6 layer MLP Dynamics Model



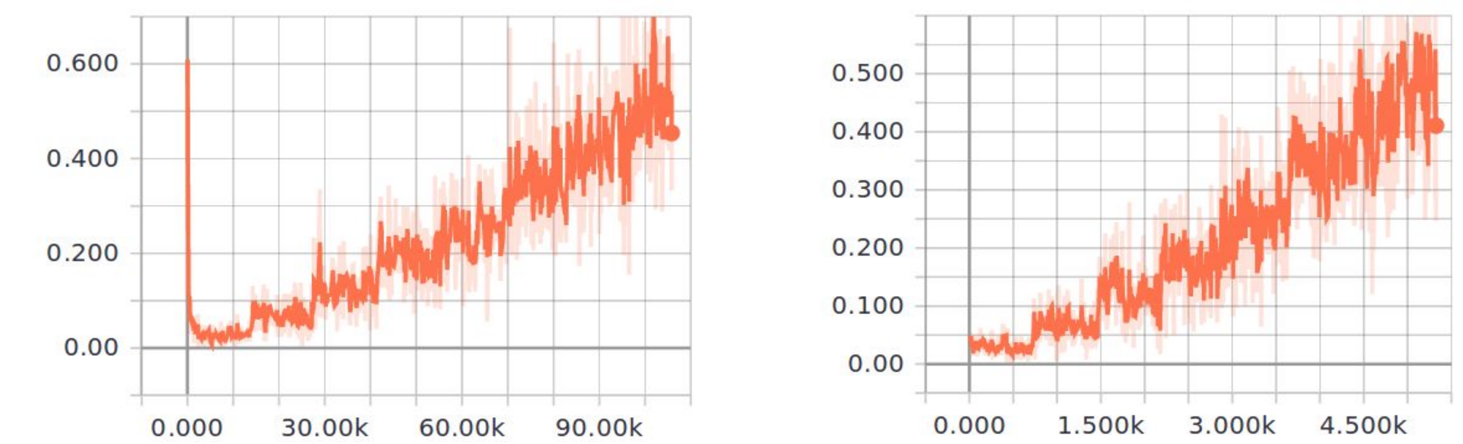
$$Loss = \log(p(image_{pred}|z_t)) + \log(p(prop_{pred}|z_t)) + Loss_{rel}$$



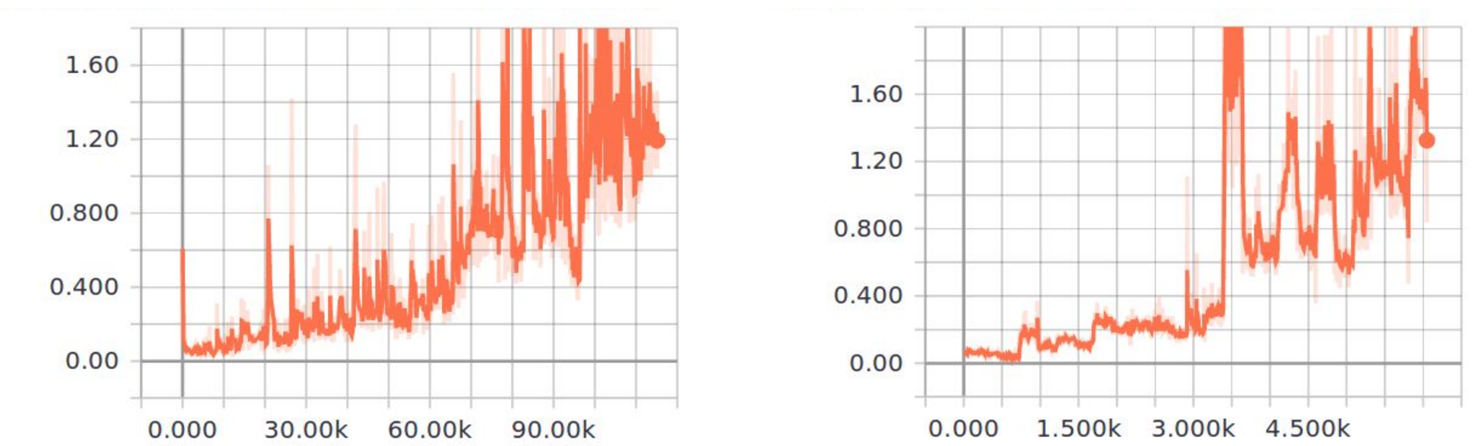
Training Details and Results:

- Optimization Algorithm: ADAM
- Curriculum: 1 -10 step prediction, 3 epochs per number of steps

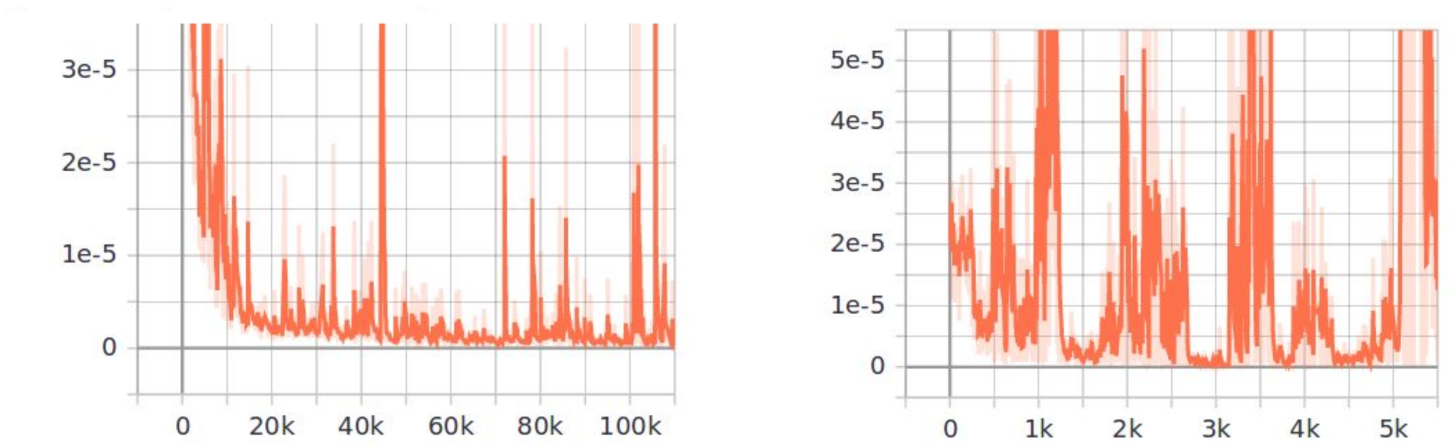
Model 1 - KL divergence Loss Term (ordering: Train, Val)



Model 2 - KL divergence Loss Term



Model 3 - Mean Square Error Dynamics Loss



Latent Space Planning Method

1. Sample a random 5 step action sequence and move away from a goal position by following that action sequence

2. Solve

$$\min_{a_{1:5}} ||z_{goal} - z_{final}||_2$$

$$s.t.\ \ z_{init} = z_{current}\ \ \ z_{t+1} = learnedmodel(z_t)$$

To solve this optimization the Cross Entropy Method was used along with the learned dynamics model.

3. Perform first action in the optimal action sequence and repeat 2.

4. The process ends when you have performed 5 actions

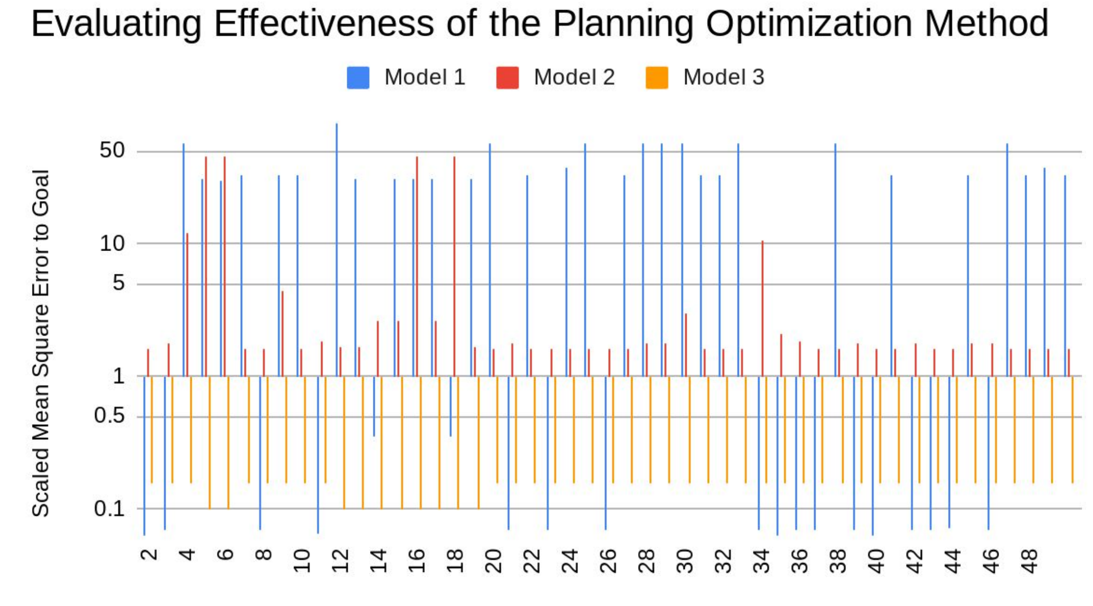
Requirements for Effective Distance Based Latent Space Planning

- Points in latent space in between z and z_goal correspond to points in the environment between x and x_goal and from which x_goal is reachable

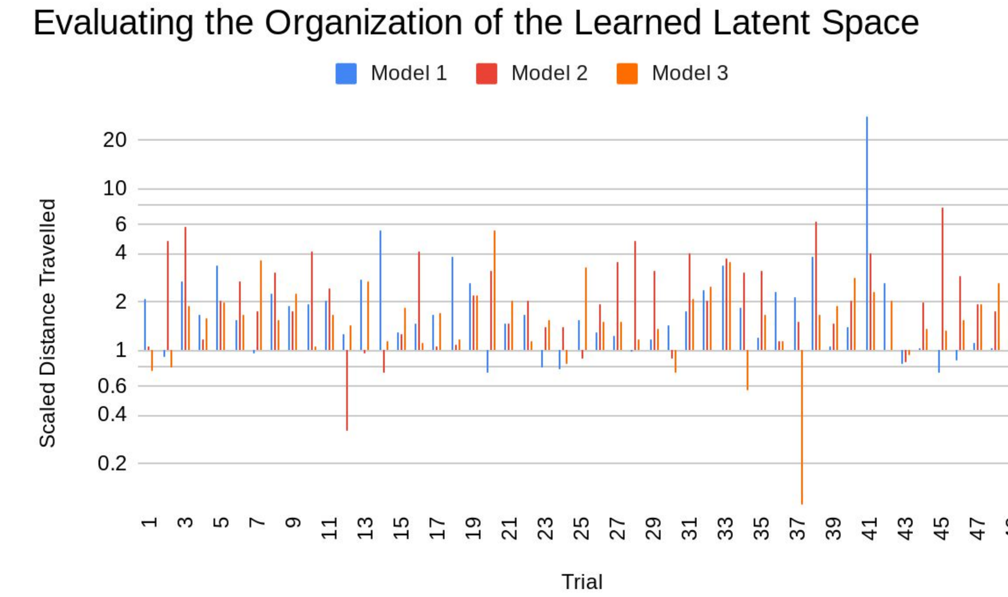
- The dynamics model can accurately predict effect of action on the environment

Latent Space Planning Results

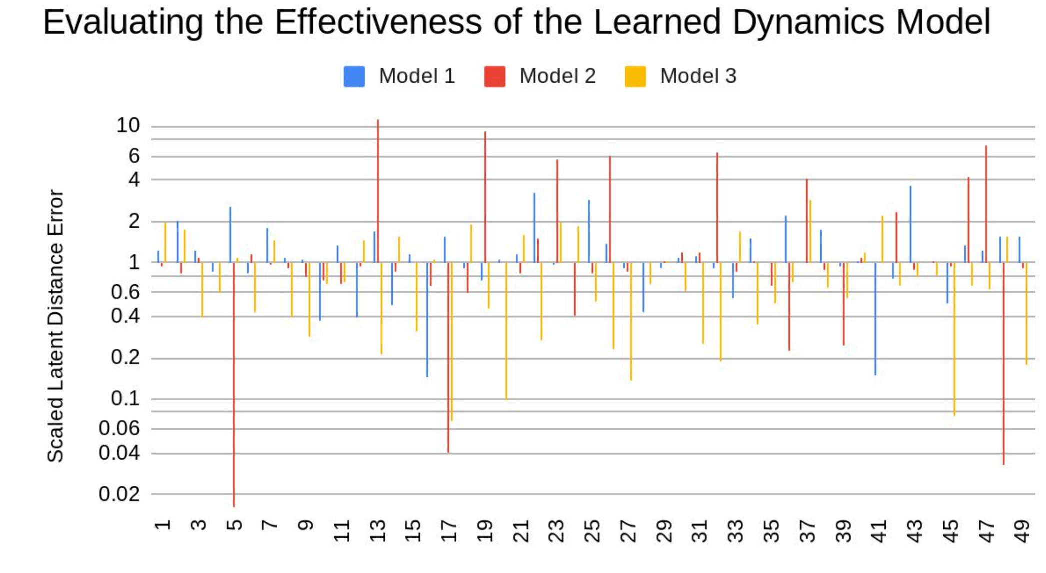
Evaluating Effectiveness of the Planning Optimization Method



Evaluating the Organization of the Learned Latent Space



Evaluating the Effectiveness of the Learned Dynamics Model



References:

1. Kingma, D. P. and Welling, M. Auto-encoding variational bayes. CoRR , abs/1312.6114, 2013. 2. Lee, M. A. Making sense of vision and touch: Self-supervised learning of multimodal representations for contact-rich tasks CoRR , abs/1810.10191, 2018. 3. Rubinstein, R. The cross-entropy method for combinatorial and continuous optimization.Methodology And Computing In Applied Probability, 1(2):127–190, Sep 1999. ISSN 1573-7713. doi: 10.1023/A:1010091220143 4. Wu, M. and Goodman, N. Multimodal generative models for scalable weakly-supervised learning. CoRR abs/1802.05335, 2018

Acknowledgement:
Thank you for the mentorship from Michelle Lee and Jeannette Bohg in the Interactive Perception and Robotic Learning Lab as Stanford