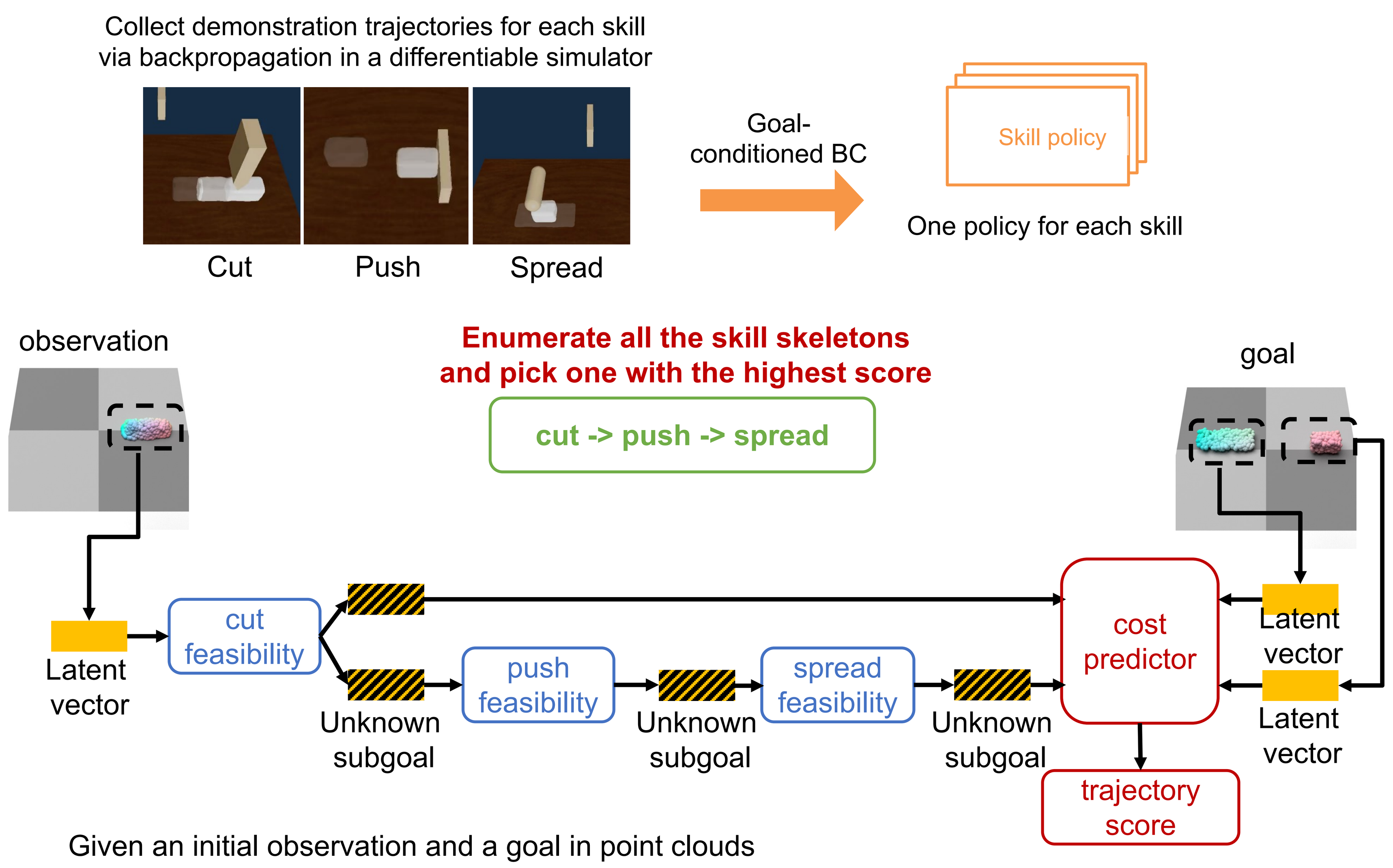
CAREER: Self-supervised Representation Learning for Deformable Object Manipulation

David Held, Carnegie Mellon University
https://r-pad.github.io/

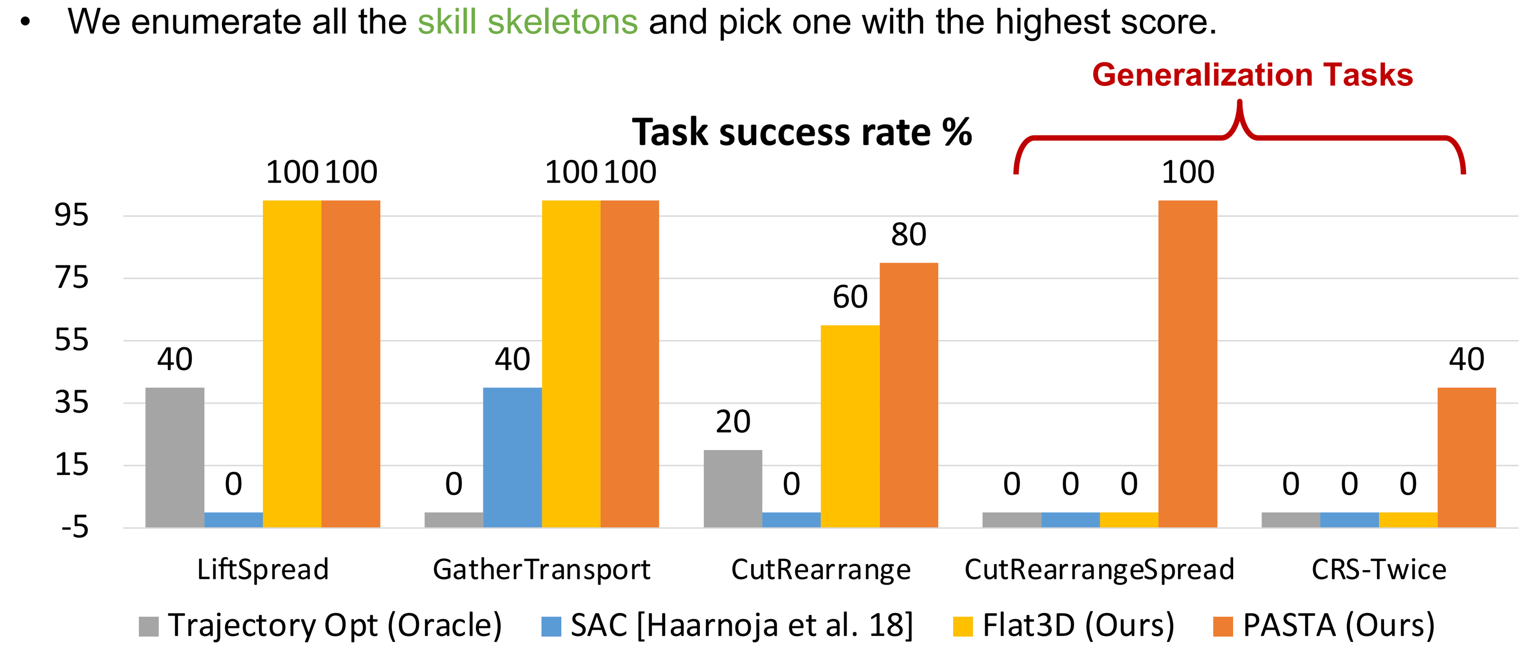
Planning with Spatial-Temporal Abstractions from Point Clouds for Deformable Object Manipulation

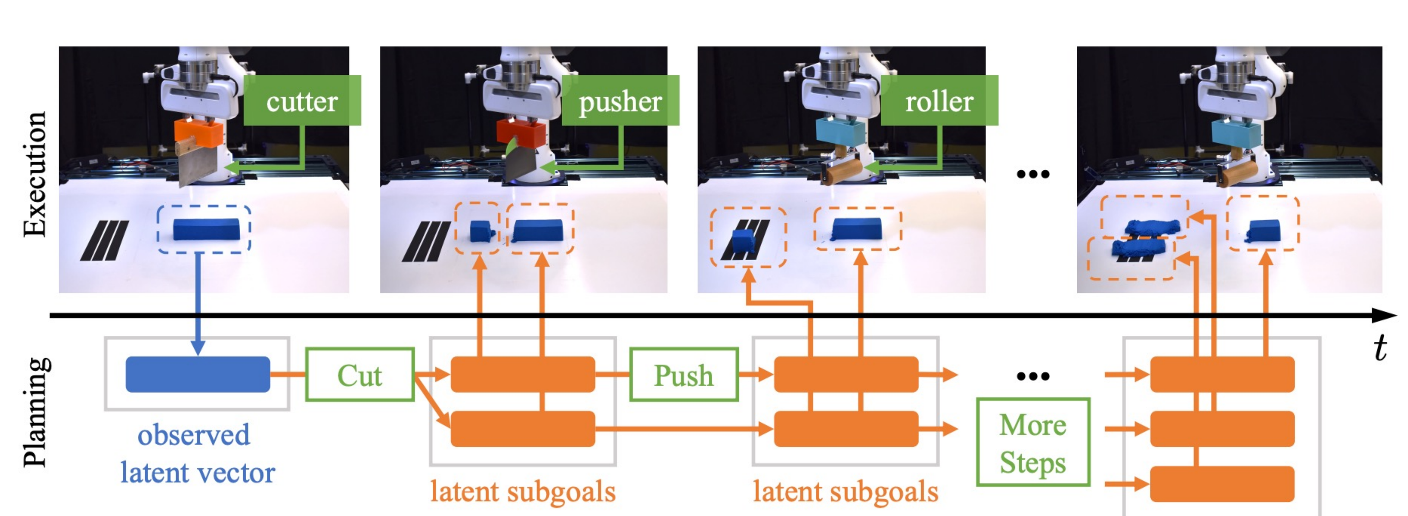
- Given a set of demonstrations of primitive skills with deformable objects, how can we compose the skills to solve long-horizon tasks from raw observations?



Given an initial observation and a goal in point clouds

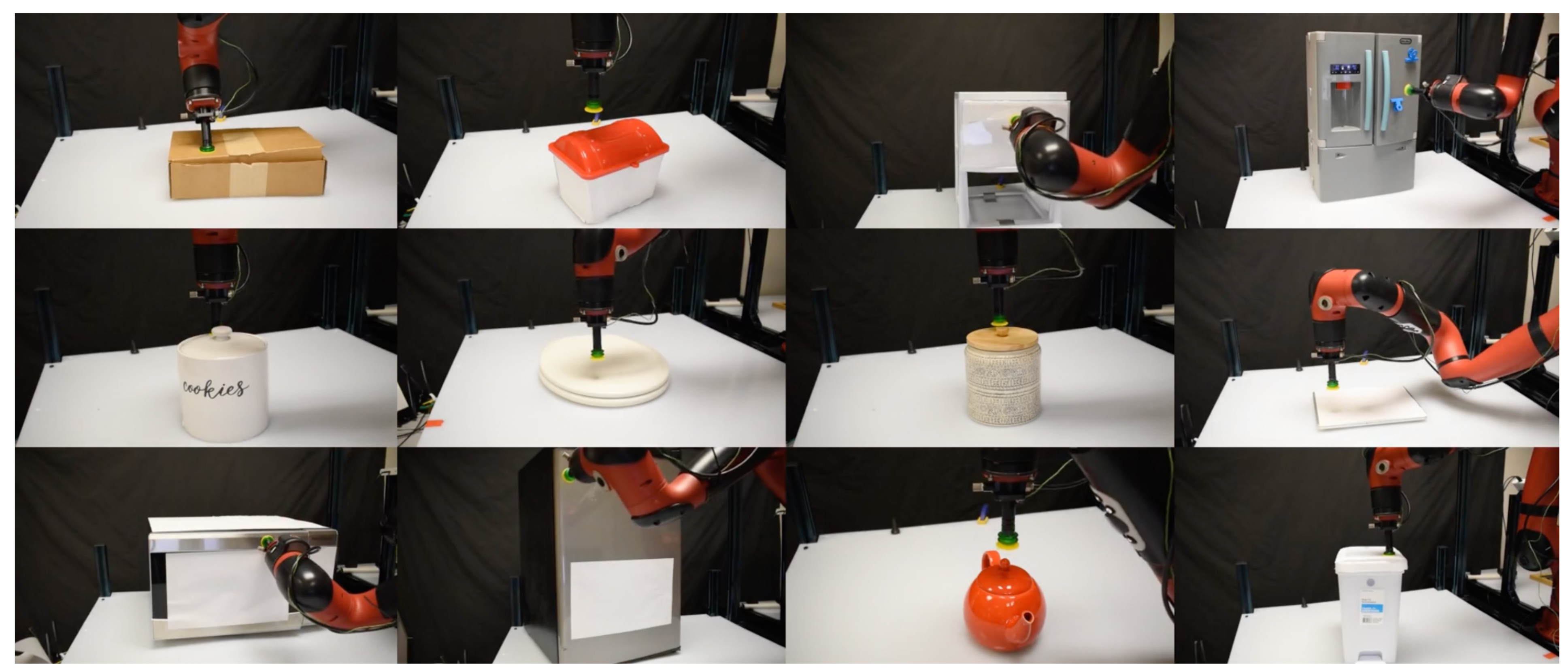
- We encode each scene into a latent set representation
- We optimize the latent subgoals with the scores predicted by the feasibility and cost predictors.
- We enumerate all the skill skeletons and pick one with the highest score.

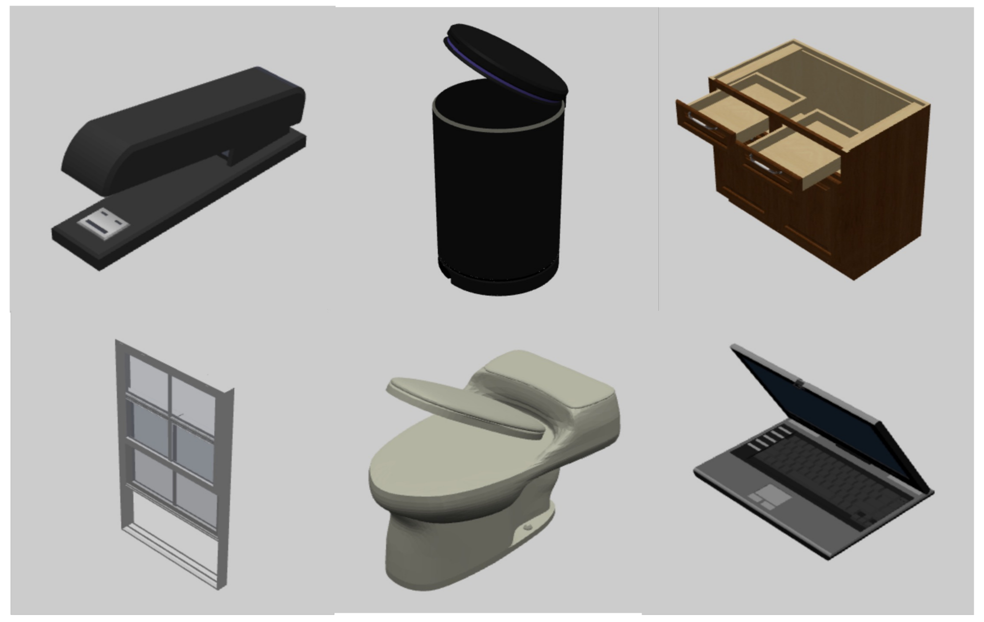




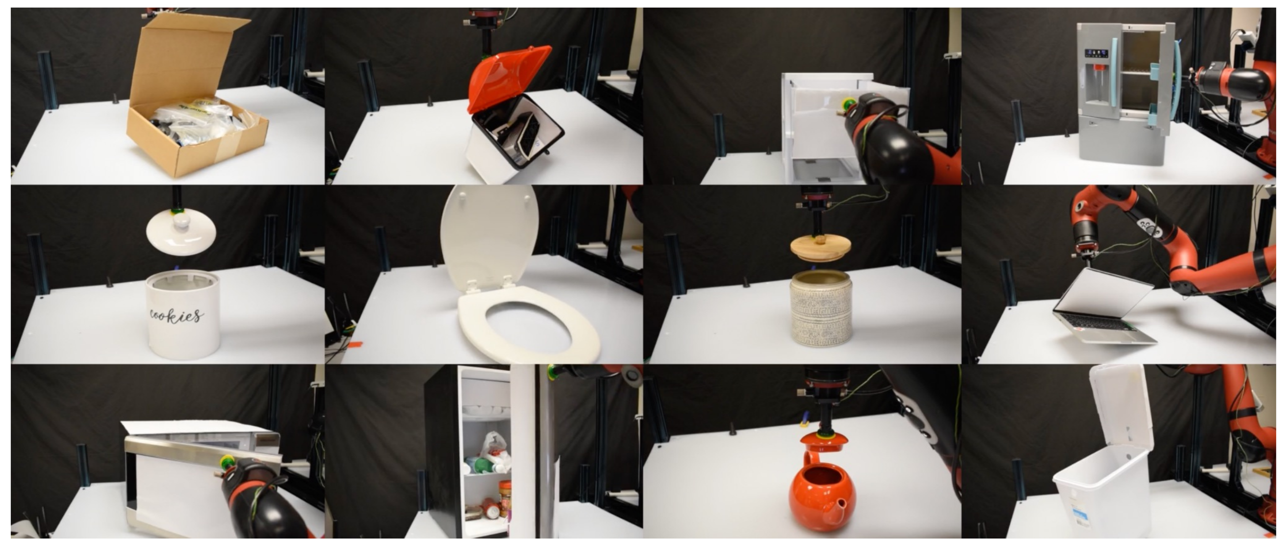
FlowBot3D: Learning 3D Articulation Flow to Manipulate Articulated Objects

- Articulated objects are everywhere (doors, drawers…).
- We want robots to be able to articulate unseen objects





Training Categories:
499 PartNet-Mobility objects from 11 categories



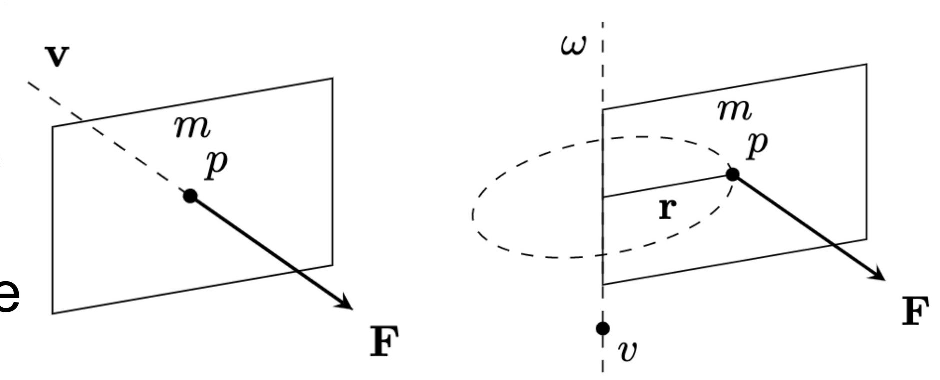
Real World Results

Previous Approaches

End2End: Directly optimizes the task objective, but hard to generalize to novel objects or poses.
Modular (Segmentation, Connectivity Estimation, Articulation Estimation): Better generalization (when it works), but many potential points of failure.

What we need to know:

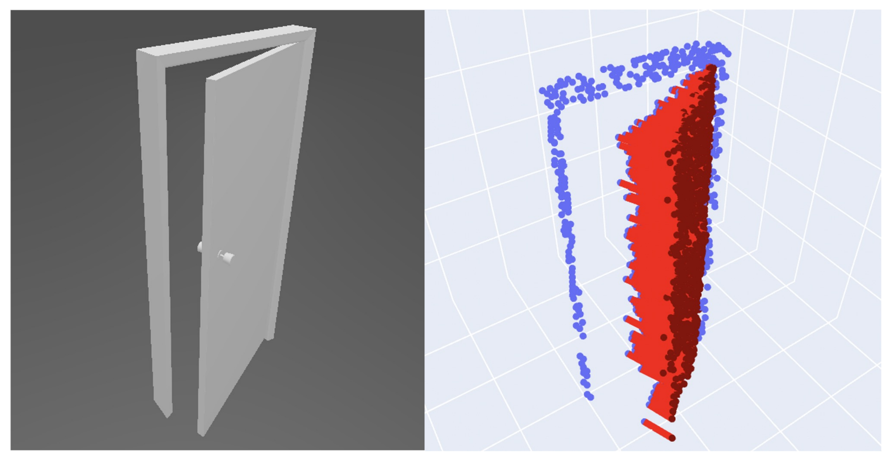
- What direction should we apply the force in?
- Which points on the object have the most leverage?

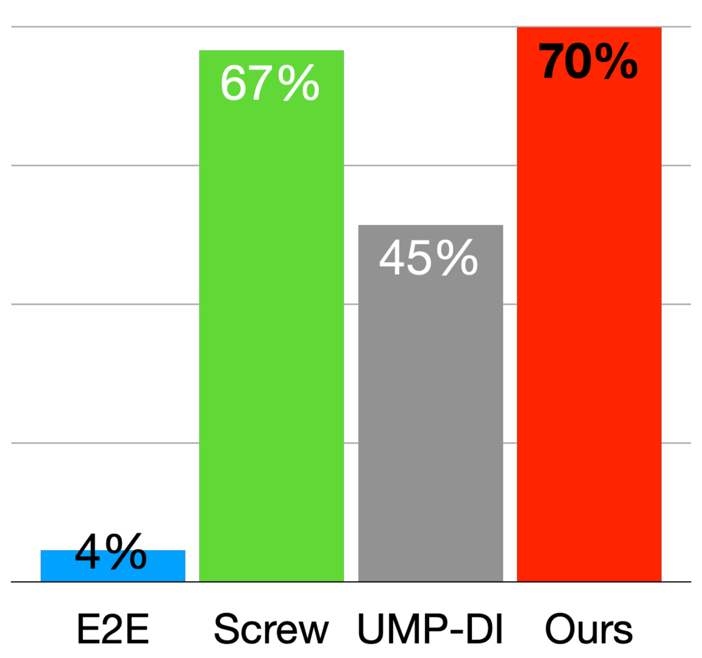


3D Articulation Flow (3DAF)

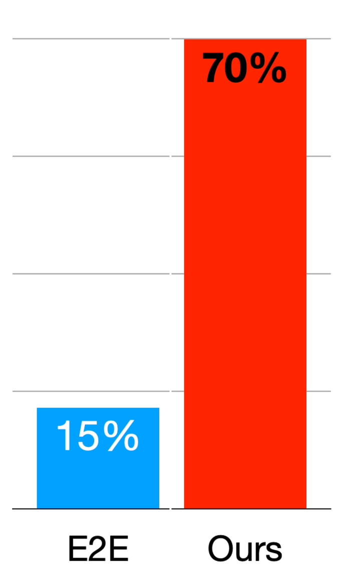
Estimate a **3D vector** for each point on an object

- **Direction**: how that point moves under **instantaneous articulation**
- **Magnitude**: relative **leverage** at that point





Unseen Categories (Sim)



Real World Results

2023 FRR & NRI Principal Investigators' Meeting
May 2-3, 2023

Award ID#: 2046491