A Baselines

For shape estimation, we compare with analytical and learning-based baselines:

Constant curvature model (Della Santina et al., 2020; Yoo et al., 2021). Constant curvature model is a common representation for the continuum deformation behavior of soft robot that parametrizes the shape with a single curvature curve (Zhang et al., 2023). Typically, the independent parameters of the state of the robot are defined by r_{curve} and θ_{curve}. Assuming a constant length, L_{curve} of the robot, we get the constraint:

$$L_{curve} = r_{curve}\theta_{curve}.$$

In typical applications, additional term ϕ_{curve} is introduced to represent the plane of bending (Katzschmann et al., 2019). We implemented this simplified representation for soft robot shape using the proposed strain model as outlined in Section 4.3 and fitting r_{curve} and θ_{curve} to the observed strains in each side of the curve. We transformed the cross-section boundary to the curve during the evaluation and measured the chamfer distance to the reference.

DeepSoRo (Wang et al., 2020). DeepSoRo architecture deploys a FoldingNet (Yang et al., 2018) decoder conditioned on visual observations to predict the current shape of a deformable body. Crucially, it is trained with chamfer distance and originally trained on partial real-world shape observations, resulting in partial point cloud reconstruction outputs without frame-to-frame correspondences. Additionally, the model directly outputs the point cloud positions in contrast to KineSoft, which learns a deformation field and produces vertex displacement with frame-to-frame correspondences. We augment DeepSoRo for evaluation by training the model on KineSoft's simulated training data and using the proposed domain alignment process.

Shape-tracking Baselines. For shape tracking and task performance evaluation we provide the results against the following: **Strain Policy**: Strain policy, based on prior works that directly use sensor readings without intermediate representations for learning manipulation policies (Sieler & Brock, 2023), uses raw sensor measurements instead of reconstructed shapes. For shape tracking evaluation, we modified the low-level controller from Section 4.4 to track reference sensor readings directly through proportional tendon actuation. For task performance evaluation, we trained a diffusion policy using the same 50 demonstrations we use for KineSoft, but with raw strain signals and wrist-mounted camera observations as input states.
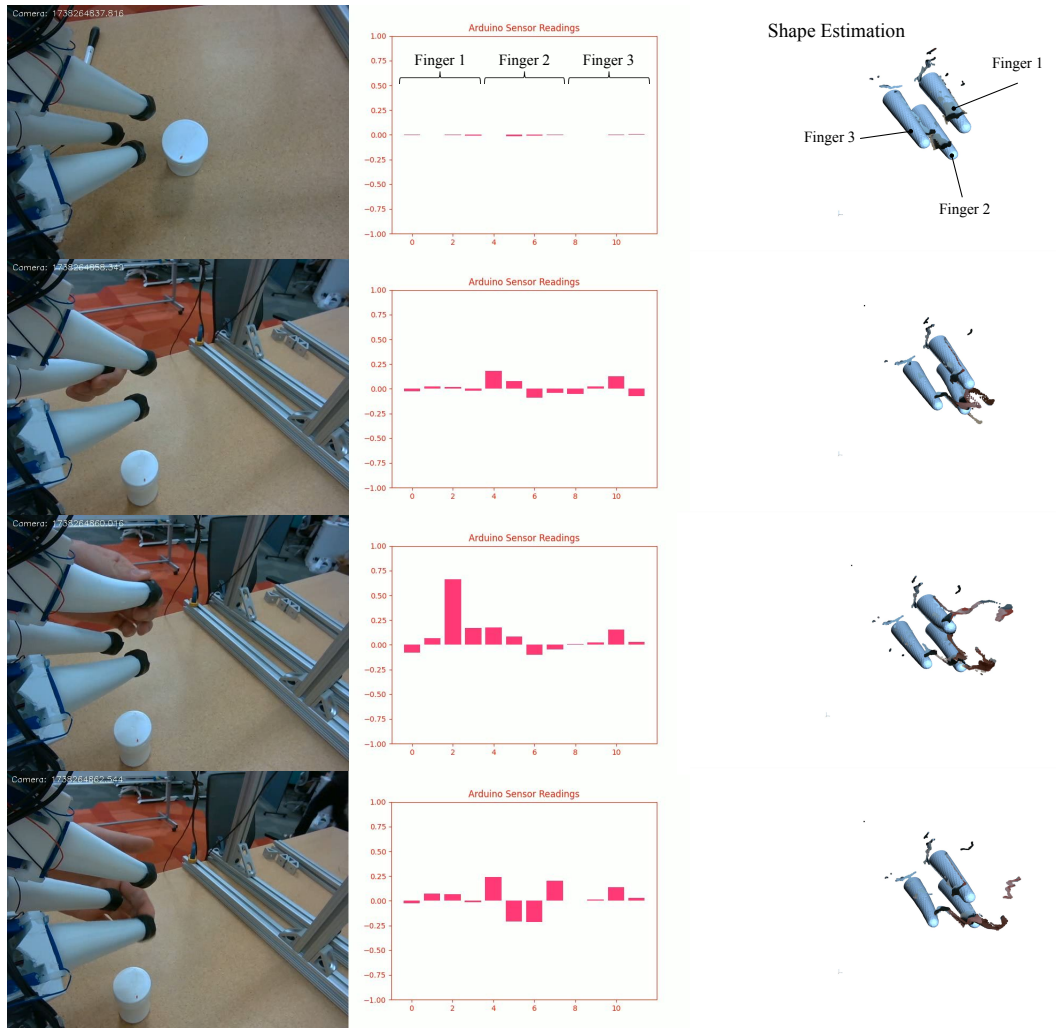
B   SENSOR SIGNALS

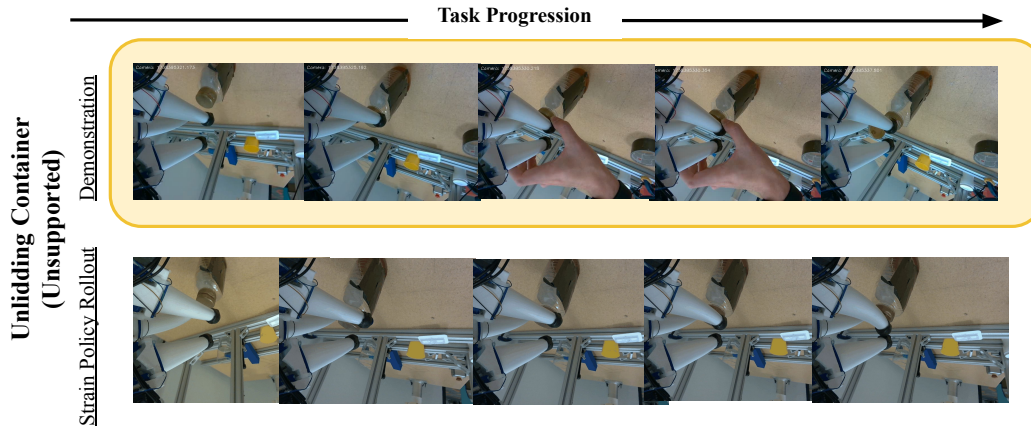Figure 6: **Sensor signals and corresponding shape estimation**

C OTHER TASK

Figure 7: **Strain tracking performance on a simple task**

Strain tracking can succeed in some tasks. Particularly with a weighted bottle opening task, where one finger could flick the lid open, the strain matching policy seemed to perform consistently.