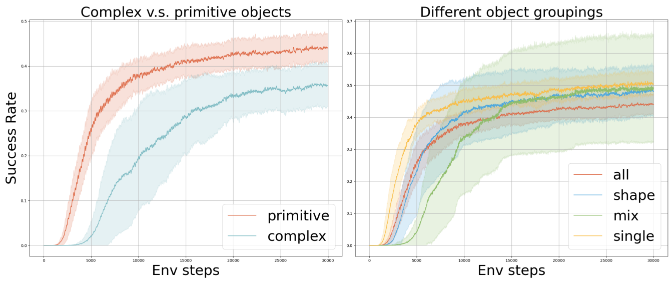

Fig. 4: **Training `grasp-and-reach` policy with different object sets.** Each curve is computed from the statistics of 10 training runs with different random seeds. Left: training with complex objects v.s. simple geometric primitive objects. Right: training with differently grouped geometric objects.

cally, we find that modeling objects as primitive geometric shapes (cylinders, cubes, and spheres) strikes a good balance between training efficiency and sim-to-real transferability. In Figure 4 (left), we compare the training curves of `grasp-and-reach` policies with primitive shapes against those with complex shapes, and the former setting is more sample efficient. More importantly, policies trained with randomized primitive shapes can also generalize to a variety of unseen objects, as shown in our video.

D. Evaluation of Reward Design

Task capabilities. With our proposed reward design principle, a wide range of long-horizon contact-rich tasks can be accomplished with pure reinforcement learning, as shown in Figure 3 and our video. The learned policies exhibit a high degree of dexterity and robustness against random force disturbances.

Effectiveness of contact-based rewards. In Figure 5, we visualize different contact behaviors that emerged from different placements or contact markers, using the `box lift` task as an example. We find that contact behaviors align closely with the contact positions specified, showing the effectiveness of using contact markers to specify contact goals.

E. Evaluation of Policy Learning

Effectiveness of task-aware hand pose initialization. In Table II, we compare the percentage of successfully trained policies for each task with and without task-aware hand pose initialization. The empirical results show that having human priors upon initialization can greatly improve the exploration efficiency of hard RL tasks.

Divide-and-conquer distillation. We evaluate our divide-and-conquer distillation approach through two ablation studies. First, we study the effect of divide-and-conquer granularity on training efficiency. Specifically, we experiment with dividing a multi-object `grasp-and-reach` task into sub-tasks that handle different numbers of objects. Starting from a total of 10 objects, we experiment with three task designs: (1) training with all objects in one policy (*all*); (2) training with three groups of similarly shaped objects in three policies (*shape*); (3) training with three groups of differently shaped objects in three policies (*mix*); (4) training with ten single-object policies (*single*). As shown in Figure 4, sample efficiency is highest for *single*, followed by *shape*, *all*, and *mix*. There is also a noticeable difference in the average success rates of each task, which could be explained as an indicator of task difficulty. Interestingly, while training with a reduced object set all reaches the same performance, training all objects in one policy shows a consistently lower performance. Second, we study the sim-to-real transfer success rate of each type of policy. Over 30 trials of each policy on an in-distribution object, we find that sim-to-real performance of the `mix` policy is the highest (90.0%), followed by `shape` (63.3%), `single` (40.0%), and `all` (23.3%). Based on the qualitative behavior of policies, we hypothesize that the low success rates of `mix` and `single` policies stem from overfitting to specific geometries, and that of `all` policy correlates with its worse performance during RL training. These results suggest that divide-and-conquer distillation helps achieve a good balance between policy training and sim-to-real transfer performance.

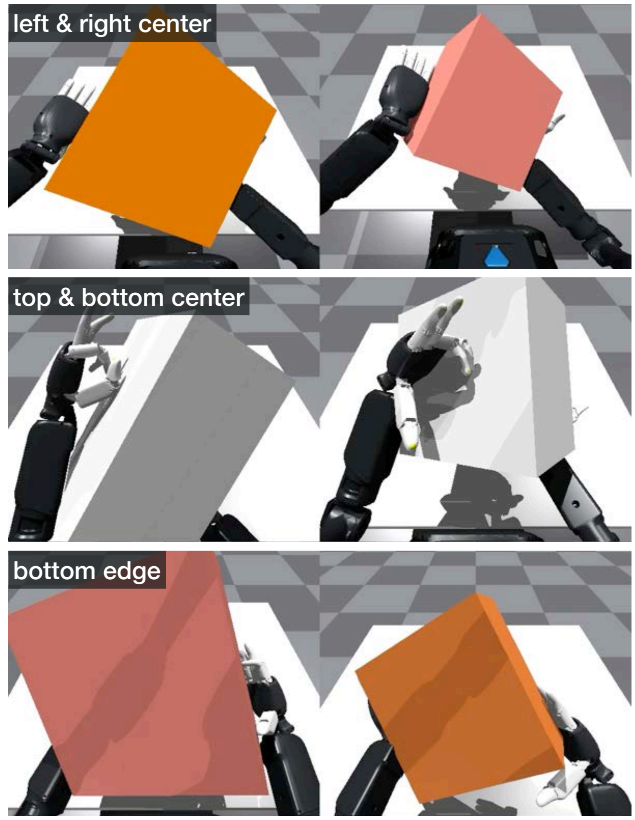

Fig. 5: **Different contact patterns emerge from different placements of contact markers.** Top: contact markers on the left and right side centers; middle: markers on the top and bottom side centers; bottom: markers on the bottom side edges.

F. Evaluation of Vision-Based Sim-to-Real Transfer

Effectiveness of mixing object representations. We investigate the effect of using different object representations and show the sim-to-real transfer comparisons in Table III. These results suggest combining dense object representation (segmented depth image) and sparse object representation (3D