# Homework 7 - Problem 1

June 14, 2021

*Part 1*

*Reward function*

My reward function is divided into 3 stages: 1. before the end effector goes near the handle, 2. when the end effector is near the handle, but the object has not reached the desired pose, 3. When the object has reached the desired pose.

In stage 1, the reward is the negative of the distance of one of the finger to the object.

In stage 2, the reward is the current object pose minus the target object pose plus a constant 2.

In stage 3, the reward is the negative of the link velocity plus the negative of the link angular velocity plus a constant 5.

I needed to add the constants $2, 5$ to stage 2 and 3 to encourage the agent to proceed from stage 1 to stage 2 and from stage 2 to stage 3.

The implementation of the compute_reward function is shown below:

```python
def compute_reward(self, action, state=None):
    info_dict = {}
    info_dict['state_info'] = {
        'gripper_pos': self.agent.robot.get_qpos()[:3].tolist(),
        'qpos': self.cabinet.get_qpos()[self.target_index_in_active_joints],
        'target_qpos': self.target_qpos,
        'link_vel': np.linalg.norm(self.target_link.get_velocity()),
        'link_ang_vel': np.linalg.norm(self.target_link.get_angular_velocity()),
    }

    si = info_dict['state_info']

    open_enough = self.cabinet.get_qpos(
    )[self.target_index_in_active_joints] >= self.target_qpos

    # If the qpos of cabinet is greater than target qpos...
```

```python
    if open_enough:

        # then we penalize the velocity.
        reward = - si['link_vel'] - si['link_ang_vel'] + 5

        return reward, info_dict

# (1, 6)
custom_obs = self.get_custom_observation()

ee_relative_coords = custom_obs['ee_relative_coords']

# 2 for the 2 fingers, 3 for relative coords
ee_relative_coords = ee_relative_coords.reshape(2, 3)

# 1 because we only use 1 finger
ee_distance_to_object = np.linalg.norm(ee_relative_coords[1])

# If the ee is too far from handle...
if ee_distance_to_object > 0.025:

    # encourage it to go near the handle.
    reward = - ee_distance_to_object

# Otherwise, try to open the door
else:

    reward = custom_obs['object_pose'] - si['target_qpos'] + 2

return reward, info_dict
```

To implement CEM, in each iteration of CEM:

- I maintain $self.popsize$ trajectories, each consisting of $self.plan\_horizon$ transitions.

- The actions in each trajectories are sampled from the current mean and std.

- For each trajectories, given the initial state and the sample action sequences, to obtain the remaining states in the trajectories, I use the function $self.predict\_next\_state\_gt$.

- For each trajectories, I obtain the sum of rewards for state-action pairs in the trajectories.

- I rank the trajectories using the sum of rewards and select $self.num_elites$ trajectories to obtain the new mean and std.

To implement MPC:

- Given a new state to plan action for, I run CEM and use the first action in the sequences of mean action as the action to take in the environment.

- After planning for action at timestep $t$, I use the resulting mean and std as the starting mean and std when replanning for action at timestep $t + 1$. When planning for action at timestep $t + 1$, the mean and std for the action at planning horizon $t + 1 + self.plan\_horizon$ are initialized to the default value.

My implementation of cem_optimize is below:

```python
def cem_optimize(self, state):
    mean = self.mean.copy()
    std = self.std.copy()
    initial_state = state.copy()

    min_std = 0.01

    for i in range(self.max_iters):

        # s (self.popsize, state size) after these 2 lines
        s = initial_state[None, :]
        s = np.repeat(s, self.popsize, axis=0)

        # ensure std is always of a certain minimum value
        std[std < min_std] = min_std

        # sample actions from mean and std
        a = np.random.normal(
            mean,
            std,
            size=(
                self.popsize,
                self.plan_horizon * self.action_dim
            )
        )

        # used to keep track of the rewards of all state-action pair
        r = np.empty((self.popsize, self.plan_horizon))

        # Roll out trajectories
        for t in range(self.plan_horizon):

            ad = self.action_dim

            actions_to_take = a[
                :,  # indexing across the population
                # selecting the t-th action in the sampled action
                t*ad:(t+1)*ad
            ]

            s, sa_r = self.predict_next_state_gt(
                s, actions_to_take
            )

            # keep track of the reward of each state-action pair
            # of the population
            r[:, t] = sa_r
```

4

```python
        # Compute the sum of reward of each sample in the population
        sor = np.sum(r, axis=1)

        # Find the elites and update mean and std
        elites = a[
            np.argsort(sor)  # ascending sort
        ][-self.num_elites:]

        mean = np.mean(elites, axis=0)
        std = np.std(elites, axis=0)

    return mean, std
```

My implementation of act is below:

```python
def act(self, state, t):
    """
    Use model predictive control to find the action give current state.

    Arguments:
        state: current state
        t: current timestep
    """
    if self.use_mpc == False:

        if t % self.plan_horizon == 0:

            self.reset()
            self.mean, self.std = self.cem_optimize(state)

        else:

            pass

        # obtain the t-th action in the mean action sequence
        ad = self.action_dim

        a_index = t % self.plan_horizon

        a = self.mean[
            a_index*ad:
            (a_index+1)*ad
        ]

        return a
    else:

        m, s = self.cem_optimize(state)

        # Pick the first action to take
        a = m[:self.action_dim]

        # Remove the first action from mean and std
        m = m[self.action_dim:]
        s = s[self.action_dim:]

        # To ensure mean and std contains plan_horizon
        # actions, initialize the mean and std
        # of the action at the end of the plan_horizon
        m = np.concatenate((m, np.zeros(self.action_dim)))
        s = np.concatenate((s, 0.5 * np.ones(self.action_dim)))
```

```python
        self.mean = m
        self.std = s

    return a
```

The complete log of a successful episode for MPC+CEM is shown below:

——step #0———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4333919584751129,
                                0.00434107007458806,
                                0.47875675559043884],
                'link_ang_vel': 1.4596707e−16,
                'link_vel': 2.7283686e−16,
                'qpos': 3.552714e−18,
                'target_qpos': 1.5607432007789612}}
reward: −0.167767 {'object_static': False, 'open_enough': False, 'success': False}
——step #1———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.40843796730041504,
                                0.013922912068665028,
                                0.4634374678134918],
                'link_ang_vel': 2.1175824e−22,
                'link_vel': 1.5825302e−16,
                'qpos': 0.0,
                'target_qpos': 1.5607432007789612}}
reward: −0.122886 {'object_static': False, 'open_enough': False, 'success': False}
——step #2———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.38347533345222473,
                                0.02194678597152233,
                                0.43906959891319275],
                'link_ang_vel': 1.4596707e−16,
                'link_vel': 2.7283686e−16,
                'qpos': 3.552714e−18,
                'target_qpos': 1.5607432007789612}}
reward: −0.075675 {'object_static': False, 'open_enough': False, 'success': False}
——step #3———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3585143983364105,
                                0.008487544022500515,
                                0.42022690176963806],
                'link_ang_vel': 2.1175824e−22,
                'link_vel': 1.5825302e−16,
                'qpos': 0.0,
                'target_qpos': 1.5607432007789612}}
reward: −0.029430 {'object_static': False, 'open_enough': False, 'success': False}
——step #4———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.36233586072921753,
                                0.010190731845796108,
                                0.4027395248413086],
                'link_ang_vel': 0.12025235,
                'link_vel': 0.010713504,

                    'qpos': −1.348791e−09,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #5———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3655101954936981,
                                  0.017474085092544556,
                                  0.39959096908569336],
                  'link_ang_vel': 0.06490808,
                  'link_vel': 0.0057827802,
                  'qpos': 1.3387445e−10,
                  'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #6———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.37051504850387573,
                                  0.02654278837144375,
                                  0.3905544877052307],
                  'link_ang_vel': 0.06850846,
                  'link_vel': 0.006103545,
                  'qpos': −3.6531445e−10,
                  'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #7———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.37616488337516785,
                                  0.025887111201882362,
                                  0.3841171860694885],
                  'link_ang_vel': 0.050057307,
                  'link_vel': 0.004459698,
                  'qpos': −3.4296266e−10,
                  'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #8———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.38320648670196533,
                                  0.01564275659620762,
                                  0.3780023157596588],
                  'link_ang_vel': 0.042500205,
                  'link_vel': 0.0037864216,
                  'qpos': −2.796327e−10,
                  'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #9———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.38774988055229187,
                                  0.015283428132534027,
                                  0.3695017397403717],

```
                    'link_ang_vel': 0.040204767,
                    'link_vel': 0.003581917,
                    'qpos': -3.1688568e-10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
----step #10----
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [-0.3912321925163269,
                                 0.012954739853739738,
                                 0.3650350868701935],
                    'link_ang_vel': 0.008793971,
                    'link_vel': 0.0007834711,
                    'qpos': -1.699979e-11,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
----step #11----
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [-0.3823668658733368,
                                 0.022024285048246384,
                                 0.36049574613571167],
                    'link_ang_vel': 0.029810788,
                    'link_vel': 0.0026558982,
                    'qpos': -1.7346198e-10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
----step #12----
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [-0.3862990140914917,
                                 0.01713927462697029,
                                 0.3525543510913849],
                    'link_ang_vel': 0.0464713,
                    'link_vel': 0.0041402145,
                    'qpos': -1.1013199e-10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
----step #13----
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [-0.38971978425979614,
                                 0.02183794416487217,
                                 0.35441866517066956],
                    'link_ang_vel': 0.93468076,
                    'link_vel': 0.08327258,
                    'qpos': 0.04001205,
                    'target_qpos': 1.5607432007789612}}
reward: 0.479269 {'object_static': False, 'open_enough': False, 'success': False}
----step #14----
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [-0.39264482259750366,
```

                    0.026041671633720398,
                    0.35663551092147827],
          'link_ang_vel': 0.7580895,
          'link_vel': 0.067539655,
          'qpos': 0.07856697,
          'target_qpos': 1.5607432007789612}}
reward: 0.517824 {'object_static': False, 'open_enough': False, 'success': False}
——step #15————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.39532727003097534,
                    0.02839614823460579,
                    0.3574695289134979],
          'link_ang_vel': 1.2240667e−09,
          'link_vel': 1.09054224e−10,
          'qpos': 0.07901151,
          'target_qpos': 1.5607432007789612}}
reward: 0.518268 {'object_static': False, 'open_enough': False, 'success': False}
——step #16————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3956676125526428,
                    0.02895551733672619,
                    0.3594735264778137],
          'link_ang_vel': 1.4598676e−16,
          'link_vel': 1.7156916e−16,
          'qpos': 0.07901151,
          'target_qpos': 1.5607432007789612}}
reward: 0.518268 {'object_static': False, 'open_enough': False, 'success': False}
——step #17————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.39883074164390564,
                    0.02846188098192215,
                    0.35777756571769714],
          'link_ang_vel': 3.297012e−09,
          'link_vel': 2.93737e−10,
          'qpos': 0.07901151,
          'target_qpos': 1.5607432007789612}}
reward: 0.518268 {'object_static': False, 'open_enough': False, 'success': False}
——step #18————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.40306714177131653,
                    0.029522379860281944,
                    0.35723817348480225],
          'link_ang_vel': 9.694773e−11,
          'link_vel': 8.637142e−12,
          'qpos': 0.07901151,
          'target_qpos': 1.5607432007789612}}
reward: 0.518268 {'object_static': False, 'open_enough': False, 'success': False}
——step #19————

11

{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.397879034280777,
                                0.026558173820376396,
                                0.35424888134002686],
                'link_ang_vel': 1.4598676e−16,
                'link_vel': 1.7156916e−16,
                'qpos': 0.07901151,
                'target_qpos': 1.5607432007789612}}
reward: 0.518268 {'object_static': False, 'open_enough': False, 'success': False}
——step #20————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3934156000614166,
                                0.03399902954697609,
                                0.35363367199897766],
                'link_ang_vel': 4.8057642e−09,
                'link_vel': 4.2815435e−10,
                'qpos': 0.07901151,
                'target_qpos': 1.5607432007789612}}
reward: 0.518268 {'object_static': False, 'open_enough': False, 'success': False}
——step #21————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.39233508706092834,
                                0.03304136171936989,
                                0.3557794690132141],
                'link_ang_vel': 0.78977054,
                'link_vel': 0.07036219,
                'qpos': 0.11373618,
                'target_qpos': 1.5607432007789612}}
reward: 0.552993 {'object_static': False, 'open_enough': False, 'success': False}
——step #22————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3961774706840515,
                                0.02697083353996277,
                                0.3567565679550171],
                'link_ang_vel': 1.3442416,
                'link_vel': 0.11976159,
                'qpos': 0.1813455,
                'target_qpos': 1.5607432007789612}}
reward: 0.620602 {'object_static': False, 'open_enough': False, 'success': False}
——step #23————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.40277212858200073,
                                0.01896682381629944,
                                0.3571683168411255],
                'link_ang_vel': 1.4771833,
                'link_vel': 0.13160557,
                'qpos': 0.25818038,
                'target_qpos': 1.5607432007789612}}

12

reward: 0.697437 {'object_static': False, 'open_enough': False, 'success': False}
——step #24————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4040735065937042,
                                0.015049010515213013,
                                0.35664352774620056],
                'link_ang_vel': 0.9387422,
                'link_vel': 0.08363444,
                'qpos': 0.30762267,
                'target_qpos': 1.5607432007789612}}
reward: 0.746879 {'object_static': False, 'open_enough': False, 'success': False}
——step #25————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4088521897792816,
                                0.0068792859092354774,
                                0.35606154799461365],
                'link_ang_vel': 0.863239,
                'link_vel': 0.0769077,
                'qpos': 0.351013,
                'target_qpos': 1.5607432007789612}}
reward: 0.790270 {'object_static': False, 'open_enough': False, 'success': False}
——step #26————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.40924394130706787,
                                0.005065068602561951,
                                0.3584011495113373],
                'link_ang_vel': 0.7438881,
                'link_vel': 0.066274434,
                'qpos': 0.3901822,
                'target_qpos': 1.5607432007789612}}
reward: 0.829439 {'object_static': False, 'open_enough': False, 'success': False}
——step #27————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4124998152256012,
                                −0.0017470518359914422,
                                0.3574787974357605],
                'link_ang_vel': 0.7101532,
                'link_vel': 0.06326892,
                'qpos': 0.4268083,
                'target_qpos': 1.5607432007789612}}
reward: 0.866065 {'object_static': False, 'open_enough': False, 'success': False}
——step #28————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.41697728633880615,
                                −0.0029373369179666042,
                                0.35886746644973755],
                'link_ang_vel': 0.8204186,
                'link_vel': 0.0730927,

                    'qpos': 0.46649706,
                    'target_qpos': 1.5607432007789612}}
reward: 0.905754 {'object_static': False, 'open_enough': False, 'success': False}
——step #29————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.420924574136734,
                                 −0.006392022594809532,
                                 0.3606373369693756],
                    'link_ang_vel': 0.68467647,
                    'link_vel': 0.06099914,
                    'qpos': 0.5002706,
                    'target_qpos': 1.5607432007789612}}
reward: 0.939527 {'object_static': False, 'open_enough': False, 'success': False}
——step #30————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.41938692331314087,
                                 −0.0023508318699896336,
                                 0.3611743152141571],
                    'link_ang_vel': 0.40224728,
                    'link_vel': 0.035836957,
                    'qpos': 0.52107877,
                    'target_qpos': 1.5607432007789612}}
reward: 0.960335 {'object_static': False, 'open_enough': False, 'success': False}
——step #31————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.42399516701698303,
                                 0.000613305892329663,
                                 0.3550734221935272],
                    'link_ang_vel': 7.586715e−09,
                    'link_vel': 6.7591427e−10,
                    'qpos': 0.52099735,
                    'target_qpos': 1.5607432007789612}}
reward: 0.960254 {'object_static': False, 'open_enough': False, 'success': False}
——step #32————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4288595914840698,
                                 0.006752100773155689,
                                 0.36125826835632324],
                    'link_ang_vel': 1.7222916e−08,
                    'link_vel': 1.5344214e−09,
                    'qpos': 0.52099735,
                    'target_qpos': 1.5607432007789612}}
reward: 0.960254 {'object_static': False, 'open_enough': False, 'success': False}
——step #33————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.42891716957092285,
                                 0.00214358823746443,
                                 0.36550241708755493],

14

```
                'link_ang_vel': 0.8717798,
                'link_vel': 0.07766864,
                'qpos': 0.56363076,
                'target_qpos': 1.5607432007789612}}
reward: 1.002887 {'object_static': False, 'open_enough': False, 'success': False}
——step #34————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4305618405342102,
                                −0.001874952227808535,
                                0.36543557047843933],
                'link_ang_vel': 0.50154567,
                'link_vel': 0.044683654,
                'qpos': 0.5933023,
                'target_qpos': 1.5607432007789612}}
reward: 1.032559 {'object_static': False, 'open_enough': False, 'success': False}
——step #35————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4317111670970917,
                                −0.007602797821164131,
                                0.3670754134654999],
                'link_ang_vel': 0.79932946,
                'link_vel': 0.07121382,
                'qpos': 0.63020504,
                'target_qpos': 1.5607432007789612}}
reward: 1.069462 {'object_static': False, 'open_enough': False, 'success': False}
——step #36————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.43168240785598755,
                                −0.01071091927587986,
                                0.3643682599067688],
                'link_ang_vel': 1.8726766e−08,
                'link_vel': 4.9952266e−08,
                'qpos': 0.64416796,
                'target_qpos': 1.5607432007789612}}
reward: 1.083425 {'object_static': False, 'open_enough': False, 'success': False}
——step #37————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4343000650405884,
                                −0.010798478499054909,
                                0.36056801676750183],
                'link_ang_vel': 5.337828e−10,
                'link_vel': 4.7555525e−11,
                'qpos': 0.64416546,
                'target_qpos': 1.5607432007789612}}
reward: 1.083422 {'object_static': False, 'open_enough': False, 'success': False}
——step #38————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.44403883814811707,
```

                                    −0.017487259581685066,
                                    0.3635149598121643],
                    'link_ang_vel': 0.44073138,
                    'link_vel': 0.03926558,
                    'qpos': 0.65527636,
                    'target_qpos': 1.5607432007789612}}
reward: 1.094533 {'object_static': False, 'open_enough': False, 'success': False}
———step #39————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.44617217779159546,
                                    −0.028980035334825516,
                                    0.36464065313339233],
                    'link_ang_vel': 0.6896379,
                    'link_vel': 0.06144118,
                    'qpos': 0.6874433,
                    'target_qpos': 1.5607432007789612}}
reward: 1.126700 {'object_static': False, 'open_enough': False, 'success': False}
———step #40————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4493746757507324,
                                    −0.03724486753344536,
                                    0.3628867268562317],
                    'link_ang_vel': 0.9332209,
                    'link_vel': 0.083142556,
                    'qpos': 0.7332426,
                    'target_qpos': 1.5607432007789612}}
reward: 1.172499 {'object_static': False, 'open_enough': False, 'success': False}
———step #41————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4472738802433014,
                                    −0.04724062606692314,
                                    0.3614638149738312],
                    'link_ang_vel': 1.1908743,
                    'link_vel': 0.10609743,
                    'qpos': 0.7904923,
                    'target_qpos': 1.5607432007789612}}
reward: 1.229749 {'object_static': False, 'open_enough': False, 'success': False}
———step #42————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.44684579968452454,
                                    −0.061734411865472794,
                                    0.3522215187549591],
                    'link_ang_vel': 1.3213649,
                    'link_vel': 0.117723234,
                    'qpos': 0.85421836,
                    'target_qpos': 1.5607432007789612}}
reward: 1.293475 {'object_static': False, 'open_enough': False, 'success': False}
———step #43————

16

{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.45029616355895996,
                                −0.07207193970680237,
                                0.34733110666275024],
                'link_ang_vel': 1.4072185,
                'link_vel': 0.12537219,
                'qpos': 0.9260087,
                'target_qpos': 1.5607432007789612}}
reward: 1.365265 {'object_static': False, 'open_enough': False, 'success': False}
——step #44————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.451069712638855,
                                −0.08395420014858246,
                                0.3449465036392212],
                'link_ang_vel': 1.3062248,
                'link_vel': 0.11637439,
                'qpos': 0.9924601,
                'target_qpos': 1.5607432007789612}}
reward: 1.431717 {'object_static': False, 'open_enough': False, 'success': False}
——step #45————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.455962598323822,
                                −0.09560275822877884,
                                0.34157368540763855],
                'link_ang_vel': 1.1998271,
                'link_vel': 0.10689514,
                'qpos': 1.0603312,
                'target_qpos': 1.5607432007789612}}
reward: 1.499588 {'object_static': False, 'open_enough': False, 'success': False}
——step #46————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.46216240525245667,
                                −0.10293576866388321,
                                0.33809372782707214],
                'link_ang_vel': 1.6796933,
                'link_vel': 0.14964761,
                'qpos': 1.1390564,
                'target_qpos': 1.5607432007789612}}
reward: 1.578313 {'object_static': False, 'open_enough': False, 'success': False}
——step #47————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4660865366458893,
                                −0.11199118942022324,
                                0.3381834626197815],
                'link_ang_vel': 1.1053423,
                'link_vel': 0.09847718,
                'qpos': 1.2003564,
                'target_qpos': 1.5607432007789612}}

17

reward: 1.639613 {'object_static': False, 'open_enough': False, 'success': False}
——step #48————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.46646204590797424,
                                −0.12133258581161499,
                                0.3427891731262207],
                'link_ang_vel': 1.4009987,
                'link_vel': 0.124818236,
                'qpos': 1.2714309,
                'target_qpos': 1.5607432007789612}}
reward: 1.710688 {'object_static': False, 'open_enough': False, 'success': False}
——step #49————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.46468856930732727,
                                −0.1311737447977066,
                                0.344889372587204],
                'link_ang_vel': 1.2108305,
                'link_vel': 0.10787542,
                'qpos': 1.3366487,
                'target_qpos': 1.5607432007789612}}
reward: 1.775905 {'object_static': False, 'open_enough': False, 'success': False}
——step #50————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.46237727999687195,
                                −0.13985848426818848,
                                0.34615853428840637],
                'link_ang_vel': 1.1708008,
                'link_vel': 0.104309075,
                'qpos': 1.3961027,
                'target_qpos': 1.5607432007789612}}
reward: 1.835359 {'object_static': False, 'open_enough': False, 'success': False}
——step #51————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.46014755964279175,
                                −0.144784078001976,
                                0.34892547130584717],
                'link_ang_vel': 0.8249616,
                'link_vel': 0.07349745,
                'qpos': 1.4418906,
                'target_qpos': 1.5607432007789612}}
reward: 1.881147 {'object_static': False, 'open_enough': False, 'success': False}
——step #52————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.45845794677734375,
                                −0.14698085188865662,
                                0.34846749901771545],
                'link_ang_vel': 0.5955829,
                'link_vel': 0.05306162,

18

                         'qpos': 1.4734281,
                         'target_qpos': 1.5607432007789612}}
reward: 1.912685 {'object_static': False, 'open_enough': False, 'success': False}
——step #53——
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.45549455285072327,
                                −0.14682011306285858,
                                0.34006327390670776],
                'link_ang_vel': 7.2991794e−17,
                'link_vel': 8.563241e−17,
                'qpos': 1.4734286,
                'target_qpos': 1.5607432007789612}}
reward: 1.912685 {'object_static': False, 'open_enough': False, 'success': False}
——step #54——
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.45332714915275574,
                                −0.1499250829219818,
                                0.34765592217445374],
                'link_ang_vel': 0.10801594,
                'link_vel': 0.0096233385,
                'qpos': 1.4757472,
                'target_qpos': 1.5607432007789612}}
reward: 1.915004 {'object_static': False, 'open_enough': False, 'success': False}
——step #55——
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.45169487595558167,
                                −0.15124483406543732,
                                0.3486633002758026],
                'link_ang_vel': 0.51916337,
                'link_vel': 0.046253234,
                'qpos': 1.5005583,
                'target_qpos': 1.5607432007789612}}
reward: 1.939815 {'object_static': False, 'open_enough': False, 'success': False}
——step #56——
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4558621644973755,
                                −0.15790477395057678,
                                0.34441012144088745],
                'link_ang_vel': 8.362101e−12,
                'link_vel': 7.450047e−13,
                'qpos': 1.4998084,
                'target_qpos': 1.5607432007789612}}
reward: 1.939065 {'object_static': False, 'open_enough': False, 'success': False}
——step #57——
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.45316818356513977,
                                −0.16291984915733337,
                                0.3491235375404358],

19

                    'link_ang_vel': 0.7684283,
                    'link_vel': 0.068460785,
                    'qpos': 1.5373433,
                    'target_qpos': 1.5607432007789612}}
reward: 1.976600 {'object_static': False, 'open_enough': False, 'success': False}
——step #58————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.44985565543174744,
                                −0.16682443022727966,
                                0.35078349709510803],
                    'link_ang_vel': 0.0030493694,
                    'link_vel': 0.0002716945,
                    'qpos': 1.5678594,
                    'target_qpos': 1.5607432007789612}}
reward: 4.996679 {'object_static': False, 'open_enough': False, 'success': False}
——step #59————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4479331076145172,
                                −0.18201661109924316,
                                0.35333487391471863],
                    'link_ang_vel': 2.8050474e−21,
                    'link_vel': 2.5111437e−22,
                    'qpos': 1.5678594,
                    'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #60————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4448243975639343,
                                −0.19221967458724976,
                                0.36167773604393005],
                    'link_ang_vel': 2.8050474e−21,
                    'link_vel': 2.5111437e−22,
                    'qpos': 1.5678594,
                    'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #61————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4501812756061554,
                                −0.1972695142030716,
                                0.35984712839126587],
                    'link_ang_vel': 2.8050474e−21,
                    'link_vel': 2.5111437e−22,
                    'qpos': 1.5678594,
                    'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #62————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.46393778920173645,

                                        −0.19918376207351685,
                                        0.35802367329597473],
                        'link_ang_vel': 2.8050474e−21,
                        'link_vel': 2.5111437e−22,
                        'qpos': 1.5678594,
                        'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #63————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.46209394931793213,
                                        −0.20754916965961456,
                                        0.34652844071388245],
                        'link_ang_vel': 2.8050474e−21,
                        'link_vel': 2.5111437e−22,
                        'qpos': 1.5678594,
                        'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #64————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.46321308612823486,
                                        −0.2262929528951645,
                                        0.3423616886138916],
                        'link_ang_vel': 2.8050474e−21,
                        'link_vel': 2.5111437e−22,
                        'qpos': 1.5678594,
                        'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #65————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4657106399536133,
                                        −0.22895357012748718,
                                        0.34509122371673584],
                        'link_ang_vel': 2.8050474e−21,
                        'link_vel': 2.5111437e−22,
                        'qpos': 1.5678594,
                        'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #66————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4571441113948822,
                                        −0.23201926052570343,
                                        0.3541877865791321],
                        'link_ang_vel': 2.8050474e−21,
                        'link_vel': 2.5111437e−22,
                        'qpos': 1.5678594,
                        'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #67————

21

{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4563668370246887,
                                −0.24831824004650116,
                                0.3465508222579956],
                'link_ang_vel': 2.8050474e−21,
                'link_vel': 2.5111437e−22,
                'qpos': 1.5678594,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #68————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4463767111301422,
                                −0.24025118350982666,
                                0.3394216001033783],
                'link_ang_vel': 2.8050474e−21,
                'link_vel': 2.5111437e−22,
                'qpos': 1.5678594,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #69————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.45834535360336304,
                                −0.2440420687198639,
                                0.3387591540813446],
                'link_ang_vel': 2.8050474e−21,
                'link_vel': 2.5111437e−22,
                'qpos': 1.5678594,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #70————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.474202036857605,
                                −0.24470460414886475,
                                0.3348648250102997],
                'link_ang_vel': 2.8050474e−21,
                'link_vel': 2.5111437e−22,
                'qpos': 1.5678594,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #71————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.48777541518211365,
                                −0.2460128664970398,
                                0.33888837695121765],
                'link_ang_vel': 2.8050474e−21,
                'link_vel': 2.5111437e−22,
                'qpos': 1.5678594,
                'target_qpos': 1.5607432007789612}}

reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #72————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.5101999044418335,
                                −0.24332600831985474,
                                0.3313961625099182],
                'link_ang_vel': 2.8050474e−21,
                'link_vel': 2.5111437e−22,
                'qpos': 1.5678594,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #73————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.5269469618797302,
                                −0.24375705420970917,
                                0.3466077148914337],
                'link_ang_vel': 2.8050474e−21,
                'link_vel': 2.5111437e−22,
                'qpos': 1.5678594,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #74————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.523462176322937,
                                −0.25699037313461304,
                                0.3395252227783203],
                'link_ang_vel': 2.8050474e−21,
                'link_vel': 2.5111437e−22,
                'qpos': 1.5678594,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #75————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.5253373384475708,
                                −0.2581287622451782,
                                0.34822943806648254],
                'link_ang_vel': 2.8050474e−21,
                'link_vel': 2.5111437e−22,
                'qpos': 1.5678594,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #76————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.5489286780357361,
                                −0.25437939167022705,
                                0.3480237126350403],
                'link_ang_vel': 2.8050474e−21,
                'link_vel': 2.5111437e−22,

```
                       'qpos': 1.5678594,
                       'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #77————
{'eval_info': {'object_static': True, 'open_enough': True, 'success': True},
  'state_info': {'gripper_pos': [−0.5530816316604614,
                                 −0.2535902261734009,
                                 0.3427658975124359],
                 'link_ang_vel': 2.8050474e−21,
                 'link_vel': 2.5111437e−22,
                 'qpos': 1.5678594,
                 'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': True, 'open_enough': True, 'success': True}
Bravo! Open the door!
```

*Part 2*

MPC+CEM can finish the episode successfully in 77 steps, while CEM-only can finish the episode successfully in 87 steps.

The complete log of a successful episode for CEM-only is shown below:

```
———step #0———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.3297604024410248,
                                0.00807458721101284,
                                0.42670556902885437],
                'link_ang_vel': 1.4596707e-16,
                'link_vel': 2.7283686e-16,
                'qpos': 3.552714e-18,
                'target_qpos': 1.5607432007789612}}
reward: -0.084652 {'object_static': False, 'open_enough': False, 'success': False}
———step #1———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.30498629808425903,
                                0.007894759997725487,
                                0.42197370529174805],
                'link_ang_vel': 2.1175824e-22,
                'link_vel': 1.5825302e-16,
                'qpos': 0.0,
                'target_qpos': 1.5607432007789612}}
reward: -0.046460 {'object_static': False, 'open_enough': False, 'success': False}
———step #2———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.2940956950187683,
                                0.009576196782290936,
                                0.41950276494026184],
                'link_ang_vel': 0.111911654,
                'link_vel': 0.009970414,
                'qpos': -9.840557e-10,
                'target_qpos': 1.5607432007789612}}
reward: -0.027147 {'object_static': False, 'open_enough': False, 'success': False}
———step #3———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.29983893036842346,
                                0.009892838075757027,
                                0.4211946427822113],
                'link_ang_vel': 0.05037017,
                'link_vel': 0.0044875722,
                'qpos': -6.264278e-10,
                'target_qpos': 1.5607432007789612}}
reward: -0.026371 {'object_static': False, 'open_enough': False, 'success': False}
———step #4———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
```

```
'state_info': {'gripper_pos': [-0.30165255069732666,
                               0.006615286692976952,
                               0.42508605122566223],
               'link_ang_vel': 0.03134744,
               'link_vel': 0.0027928012,
               'qpos': -1.1233769e-10,
               'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #5———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.3113330006599426,
                                -0.0038103139959275723,
                                0.42818430066108704],
                'link_ang_vel': 0.07029165,
                'link_vel': 0.0062624128,
                'qpos': -5.221197e-10,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #6———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.3228222727775574,
                                0.0032427653204649687,
                                0.42520201206207275],
                'link_ang_vel': 0.045611426,
                'link_vel': 0.004063606,
                'qpos': -4.7741616e-10,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #7———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.33148840069770813,
                                0.006434113252907991,
                                0.4254186749458313],
                'link_ang_vel': 0.05831935,
                'link_vel': 0.0051957783,
                'qpos': -4.0291034e-10,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #8———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.33404460549354553,
                                0.008009504526853561,
                                0.42237699031829834],
                'link_ang_vel': 0.039468963,
                'link_vel': 0.0035163627,
                'qpos': -3.3585518e-10,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
```

```
——step #9————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3417244851589203,
                                0.005997239612042904,
                                0.41911691427230835],
                'link_ang_vel': 0.053387824,
                'link_vel': 0.0047564195,
                'qpos': −3.2095393e−10,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #10————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3496636152267456,
                                −0.002420638222247362,
                                0.41684380173683167],
                'link_ang_vel': 0.034648847,
                'link_vel': 0.0030869294,
                'qpos': −3.2840447e−10,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #11————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.35248640179634094,
                                −0.0035890371073037386,
                                0.4155389070510864],
                'link_ang_vel': 0.03783995,
                'link_vel': 0.0033712315,
                'qpos': −3.656573e−10,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #12————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.35981565713882446,
                                0.0005071972846053541,
                                0.4124361574649811],
                'link_ang_vel': 0.05131609,
                'link_vel': 0.0045718453,
                'qpos': −5.370207e−10,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #13————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.36435970664024353,
                                −0.00554056977853179,
                                0.40950489044189453],
                'link_ang_vel': 0.03476418,
                'link_vel': 0.003097205,
                'qpos': −1.868434e−10,
```

                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #14————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3686712384223938,
                                0.00067782029509554437,
                                0.40400391817092896],
                'link_ang_vel': 0.051809732,
                'link_vel': 0.004615825,
                'qpos': −4.0291026e−10,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #15————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.36618348956108093,
                                −0.0006624605739489198,
                                0.40925341844558716],
                'link_ang_vel': 0.02209502,
                'link_vel': 0.0019684865,
                'qpos': −1.88706e−10,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #16————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.37021106481552124,
                                −0.0010063264053314924,
                                0.4074035882949829],
                'link_ang_vel': 0.033033293,
                'link_vel': 0.002942997,
                'qpos': −2.4458544e−10,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #17————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.36792173981666565,
                                −0.00056750129442662,
                                0.3997523784637451],
                'link_ang_vel': 0.03394489,
                'link_vel': 0.0030242128,
                'qpos': 1.991017e−11,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #18————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.36625269055366516,
                                −0.004884784109890461,
                                0.39220771193504333],
                'link_ang_vel': 0.028152926,

28

                    'link_vel': 0.002508196,
                    'qpos': −2.3713487e−10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
———step #19————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.36999037861824036,
                                 −0.009985980577766895,
                                 0.38942283391952515],
                    'link_ang_vel': 0.041865256,
                    'link_vel': 0.0037298538,
                    'qpos': −1.2537617e−10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
———step #20————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.36952826380729675,
                                 −0.007571469526737928,
                                 0.3853529095649719],
                    'link_ang_vel': 0.024534604,
                    'link_vel': 0.0021858334,
                    'qpos': −6.577153e−11,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
———step #21————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3750201165676117,
                                 −0.007300466299057007,
                                 0.38493361515045166],
                    'link_ang_vel': 0.029712629,
                    'link_vel': 0.0026471529,
                    'qpos': −3.8242126e−10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
———step #22————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3795449435710907,
                                 −0.005559755489230156,
                                 0.3757447600364685],
                    'link_ang_vel': 0.010749782,
                    'link_vel': 0.00095771794,
                    'qpos': −6.43746e−11,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
———step #23————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.38194897770881653,
                                 −0.005217096302658319,

29

```
                                    0.3761802017688751],
                    'link_ang_vel': 5.4513496e-09,
                    'link_vel': 4.856709e-10,
                    'qpos': 2.3123867e-10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
————step #24————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [-0.38210245966911316,
                                    -0.009183797053992748,
                                    0.36572256684303284],
                    'link_ang_vel': 0.019474136,
                    'link_vel': 0.0017349868,
                    'qpos': 2.605311e-10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
————step #25————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [-0.3812759518623352,
                                    -0.012767819687724113,
                                    0.3611113131046295],
                    'link_ang_vel': 0.0140760355,
                    'link_vel': 0.00125406,
                    'qpos': -1.827785e-10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
————step #26————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [-0.3812141716480255,
                                    -0.014735331758856773,
                                    0.3613266050815582],
                    'link_ang_vel': 1.0005667e-09,
                    'link_vel': 8.9142436e-11,
                    'qpos': 2.0566572e-10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
————step #27————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [-0.3840304911136627,
                                    -0.01972980797290802,
                                    0.36147674918174744],
                    'link_ang_vel': 2.3894737e-08,
                    'link_vel': 2.1288264e-09,
                    'qpos': 0.00011485521,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439372 {'object_static': True, 'open_enough': False, 'success': False}
————step #28————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
```

```
            'state_info': {'gripper_pos': [−0.38863617181777954,
                                          −0.01821545697748661,
                                          0.36050334572792053],
                  'link_ang_vel': 0.18052171,
                  'link_vel': 0.01608301,
                  'qpos': 0.008177574,
                  'target_qpos': 1.5607432007789612}}
reward: 0.447434 {'object_static': True, 'open_enough': False, 'success': False}
——step #29————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.38381627202033997,
                                          −0.016755791381001472,
                                          0.3594984710216522],
                  'link_ang_vel': 1.7090809e−09,
                  'link_vel': 1.5226513e−10,
                  'qpos': 0.008175562,
                  'target_qpos': 1.5607432007789612}}
reward: 0.447432 {'object_static': True, 'open_enough': False, 'success': False}
——step #30————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3845152258872986,
                                          −0.011076664552092552,
                                          0.3571421205997467],
                  'link_ang_vel': 3.497207e−09,
                  'link_vel': 3.1157252e−10,
                  'qpos': 0.008170114,
                  'target_qpos': 1.5607432007789612}}
reward: 0.447427 {'object_static': True, 'open_enough': False, 'success': False}
——step #31————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.38322362303733826,
                                          −0.01197643019258976,
                                          0.35993629693984985],
                  'link_ang_vel': 0.0072432943,
                  'link_vel': 0.0006453184,
                  'qpos': −2.2933818e−11,
                  'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #32————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.38402441143989563,
                                          −0.00857670046389103,
                                          0.3566074073314667],
                  'link_ang_vel': 0.004366783,
                  'link_vel': 0.00038904476,
                  'qpos': −6.204937e−11,
                  'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
```

——step #33————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.385475218296051,
                                −0.00894802063703537,
                                0.35100436210632324],
                'link_ang_vel': 0.0059922487,
                'link_vel': 0.0005338606,
                'qpos': −2.6193445e−11,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #34————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3856984078884125,
                                −0.008905656635761261,
                                0.35033005475997925],
                'link_ang_vel': 0.0057874895,
                'link_vel': 0.0005156182,
                'qpos': −9.208455e−11,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #35————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.38938793540000916,
                                −0.011994168162345886,
                                0.3548550307750702],
                'link_ang_vel': 4.860368e−19,
                'link_vel': 4.3301907e−20,
                'qpos': 0.009108744,
                'target_qpos': 1.5607432007789612}}
reward: −0.036048 {'object_static': True, 'open_enough': False, 'success': False}
——step #36————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3937744200229645,
                                −0.0056269727647304535,
                                0.34525755047798157],
                'link_ang_vel': 0.03006845,
                'link_vel': 0.0026788537,
                'qpos': −1.8998983e−10,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #37————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3987765312194824,
                                −0.004909050650894642,
                                0.349730908870697],
                'link_ang_vel': 0.014665694,
                'link_vel': 0.0013065937,
                'qpos': −2.3865415e−10,

```
                              'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
----step #38----
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.3954520523548126,
                                -0.003924610558897257,
                                0.3520650267601013],
                'link_ang_vel': 0.009310101,
                'link_vel': 0.0008294541,
                'qpos': -9.2145486e-11,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
----step #39----
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.3991454243659973,
                                -0.002943413332104683,
                                0.3554668128490448],
                'link_ang_vel': 0.66509163,
                'link_vel': 0.05925429,
                'qpos': 0.029322093,
                'target_qpos': 1.5607432007789612}}
reward: 0.468579 {'object_static': False, 'open_enough': False, 'success': False}
----step #40----
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.39610177278518677,
                                0.000854438345413655,
                                0.35726943612098694],
                'link_ang_vel': 0.09832214,
                'link_vel': 0.008759699,
                'qpos': 0.025376393,
                'target_qpos': 1.5607432007789612}}
reward: 0.464633 {'object_static': False, 'open_enough': False, 'success': False}
----step #41----
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.39397501945495605,
                                -0.005061100237071514,
                                0.3627089858055115],
                'link_ang_vel': 6.28428e-08,
                'link_vel': 1.5700763e-07,
                'qpos': 0.025898535,
                'target_qpos': 1.5607432007789612}}
reward: 0.465155 {'object_static': False, 'open_enough': False, 'success': False}
----step #42----
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.39520448446273804,
                                -0.0007462164503522217,
                                0.3703314960002899],
                'link_ang_vel': 0.6659088,
```

```
                    'link_vel': 0.0593271,
                    'qpos': 0.056479618,
                    'target_qpos': 1.5607432007789612}}
reward: 0.495736 {'object_static': False, 'open_enough': False, 'success': False}
——step #43————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.39873144030570984,
                                  0.005645228084176779,
                                  0.37163013219833374],
                    'link_ang_vel': 0.35332257,
                    'link_vel': 0.031478163,
                    'qpos': 0.07556574,
                    'target_qpos': 1.5607432007789612}}
reward: 0.514822 {'object_static': False, 'open_enough': False, 'success': False}
——step #44————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3942004144191742,
                                  0.007617031689733267,
                                  0.37706777453422546],
                    'link_ang_vel': 0.032917045,
                    'link_vel': 0.00293264,
                    'qpos': 0.07515952,
                    'target_qpos': 1.5607432007789612}}
reward: 0.514416 {'object_static': False, 'open_enough': False, 'success': False}
——step #45————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.39119988679885864,
                                  −0.005364274606108665,
                                  0.38293254375457764],
                    'link_ang_vel': 0.12708466,
                    'link_vel': 0.011322207,
                    'qpos': 0.07715501,
                    'target_qpos': 1.5607432007789612}}
reward: 0.516412 {'object_static': False, 'open_enough': False, 'success': False}
——step #46————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.39235445857048035,
                                  −0.006412900052964687,
                                  0.38997119665145874],
                    'link_ang_vel': 0.46245772,
                    'link_vel': 0.041201223,
                    'qpos': 0.09766477,
                    'target_qpos': 1.5607432007789612}}
reward: 0.536922 {'object_static': False, 'open_enough': False, 'success': False}
——step #47————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3896022439002991,
                                  −0.010565344244241714,
```

                                0.39245185256004333],
                    'link_ang_vel': 1.7106511e-09,
                    'link_vel': 1.5240496e-10,
                    'qpos': 0.09776256,
                    'target_qpos': 1.5607432007789612}}
reward: 0.537019 {'object_static': False, 'open_enough': False, 'success': False}
——step #48————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.38538384437561035,
                                −0.007957221940159798,
                                0.394395649433136],
                    'link_ang_vel': 4.7412634e-09,
                    'link_vel': 4.224077e-10,
                    'qpos': 0.09752741,
                    'target_qpos': 1.5607432007789612}}
reward: 0.536784 {'object_static': False, 'open_enough': False, 'success': False}
——step #49————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.37985286116600037,
                                −0.008733180351555347,
                                0.40123075246810913],
                    'link_ang_vel': 0.71444833,
                    'link_vel': 0.06365158,
                    'qpos': 0.11700557,
                    'target_qpos': 1.5607432007789612}}
reward: 0.556262 {'object_static': False, 'open_enough': False, 'success': False}
——step #50————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3764008581638336,
                                −0.010518273338675499,
                                0.40272257561683655],
                    'link_ang_vel': 0.21375687,
                    'link_vel': 0.019043997,
                    'qpos': 0.12807126,
                    'target_qpos': 1.5607432007789612}}
reward: 0.567328 {'object_static': False, 'open_enough': False, 'success': False}
——step #51————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.37234240770339966,
                                −0.007572307251393795,
                                0.39857158064842224],
                    'link_ang_vel': 0.36344966,
                    'link_vel': 0.03238041,
                    'qpos': 0.14526352,
                    'target_qpos': 1.5607432007789612}}
reward: 0.584520 {'object_static': False, 'open_enough': False, 'success': False}
——step #52————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},

'state_info': {'gripper_pos': [−0.3712279200553894,
                                −0.011524390429258347,
                                0.40567266941070557],
                'link_ang_vel': 1.2235305,
                'link_vel': 0.10900694,
                'qpos': 0.19976223,
                'target_qpos': 1.5607432007789612}}
reward: 0.639019 {'object_static': False, 'open_enough': False, 'success': False}
——step #53————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3706924617290497,
                                −0.019272247329354286,
                                0.414110004901886],
                'link_ang_vel': 2.278898,
                'link_vel': 0.20303325,
                'qpos': 0.3065414,
                'target_qpos': 1.5607432007789612}}
reward: 0.745798 {'object_static': False, 'open_enough': False, 'success': False}
——step #54————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3698819875717163,
                                −0.021491222083568573,
                                0.41579458117485046],
                'link_ang_vel': 0.61014605,
                'link_vel': 0.054359064,
                'qpos': 0.34580165,
                'target_qpos': 1.5607432007789612}}
reward: 0.785058 {'object_static': False, 'open_enough': False, 'success': False}
——step #55————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.37109625339508057,
                                −0.022037534043192863,
                                0.4184006452560425],
                'link_ang_vel': 1.3327429,
                'link_vel': 0.11873694,
                'qpos': 0.40901193,
                'target_qpos': 1.5607432007789612}}
reward: 0.848269 {'object_static': False, 'open_enough': False, 'success': False}
——step #56————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3660263419151306,
                                −0.027010973542928696,
                                0.42215391993522644],
                'link_ang_vel': 2.1857884,
                'link_vel': 0.19473772,
                'qpos': 0.5122254,
                'target_qpos': 1.5607432007789612}}
reward: 0.951482 {'object_static': False, 'open_enough': False, 'success': False}

——step #57————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3588567078113556,
                                −0.03412046656012535,
                                0.42573434114456177],
                'link_ang_vel': 2.1745381,
                'link_vel': 0.19373524,
                'qpos': 0.6061363,
                'target_qpos': 1.5607432007789612}}
reward: 1.045393 {'object_static': False, 'open_enough': False, 'success': False}
——step #58————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.36133190989494324,
                                −0.03828132897615433,
                                0.42688190937042236],
                'link_ang_vel': 1.7212249,
                'link_vel': 0.15334798,
                'qpos': 0.6906084,
                'target_qpos': 1.5607432007789612}}
reward: 1.129865 {'object_static': False, 'open_enough': False, 'success': False}
——step #59————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3650016188621521,
                                −0.036790817975997925,
                                0.42732980847358704],
                'link_ang_vel': 0.9952153,
                'link_vel': 0.08866574,
                'qpos': 0.7444401,
                'target_qpos': 1.5607432007789612}}
reward: 1.183697 {'object_static': False, 'open_enough': False, 'success': False}
——step #60————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.37071922421455383,
                                −0.048292748630046844,
                                0.4280475676059723],
                'link_ang_vel': 2.9056618,
                'link_vel': 0.2588753,
                'qpos': 0.8710193,
                'target_qpos': 1.5607432007789612}}
reward: 1.310276 {'object_static': False, 'open_enough': False, 'success': False}
——step #61————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.36677926778793335,
                                −0.05977719649672508,
                                0.4307211935520172],
                'link_ang_vel': 2.5718396,
                'link_vel': 0.22913373,
                'qpos': 0.9963977,

```
                              'target_qpos': 1.5607432007789612}}
reward: 1.435654 {'object_static': False, 'open_enough': False, 'success': False}
——step #62————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.37129855155944824,
                              −0.06942512094974518,
                              0.4297759532928467],
               'link_ang_vel': 1.6347345,
               'link_vel': 0.14564228,
               'qpos': 1.0968397,
               'target_qpos': 1.5607432007789612}}
reward: 1.536096 {'object_static': False, 'open_enough': False, 'success': False}
——step #63————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.36858731508255005,
                              −0.08397156000137329,
                              0.4295695722103119],
               'link_ang_vel': 1.3473449,
               'link_vel': 0.12003796,
               'qpos': 1.1528633,
               'target_qpos': 1.5607432007789612}}
reward: 1.592120 {'object_static': False, 'open_enough': False, 'success': False}
——step #64————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.36868470907211304,
                              −0.09360865503549576,
                              0.4306042194366455],
               'link_ang_vel': 2.587308,
               'link_vel': 0.2305118,
               'qpos': 1.2775929,
               'target_qpos': 1.5607432007789612}}
reward: 1.716850 {'object_static': False, 'open_enough': False, 'success': False}
——step #65————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.36462661623954773,
                              −0.10373299568891525,
                              0.43163326382637024],
               'link_ang_vel': 2.4458687,
               'link_vel': 0.21790947,
               'qpos': 1.3985794,
               'target_qpos': 1.5607432007789612}}
reward: 1.837836 {'object_static': False, 'open_enough': False, 'success': False}
——step #66————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3628256618976593,
                              −0.1184546947479248,
                              0.432382732629776],
               'link_ang_vel': 3.0404623,
```

```
                  'link_vel': 0.27088496,
                  'qpos': 1.5325311,
                  'target_qpos': 1.5607432007789612}}
reward: 1.971788 {'object_static': False, 'open_enough': False, 'success': False}
——step #67————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.3604365289211273,
                                -0.12643444538116455,
                                0.4299934208393097],
                  'link_ang_vel': 0.9891985,
                  'link_vel': 0.088129684,
                  'qpos': 1.6102495,
                  'target_qpos': 1.5607432007789612}}
reward: 3.922672 {'object_static': False, 'open_enough': False, 'success': False}
——step #68————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.3503940999507904,
                                -0.13696768879890442,
                                0.45132070779800415],
                  'link_ang_vel': 1.1369268e-19,
                  'link_vel': 1.0129041e-20,
                  'qpos': 1.6398993,
                  'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #69————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.34944772720336914,
                                -0.15030154585838318,
                                0.47123226523399353],
                  'link_ang_vel': 1.1369268e-19,
                  'link_vel': 1.0129041e-20,
                  'qpos': 1.6398993,
                  'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #70————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.3467763364315033,
                                -0.15166078507900238,
                                0.47709813714027405],
                  'link_ang_vel': 1.1369268e-19,
                  'link_vel': 1.0129041e-20,
                  'qpos': 1.6398993,
                  'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #71————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.3473638594150543,
                                -0.15262877941131592,
```

```
                                0.48466816544532776],
                    'link_ang_vel': 1.1369268e−19,
                    'link_vel': 1.0129041e−20,
                    'qpos': 1.6398993,
                    'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #72————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3425079882144928,
                                 −0.1443474292755127,
                                 0.4795086681842804],
                    'link_ang_vel': 1.1369268e−19,
                    'link_vel': 1.0129041e−20,
                    'qpos': 1.6398993,
                    'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #73————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3425118327140808,
                                 −0.1450822651386261,
                                 0.4804023802280426],
                    'link_ang_vel': 1.1369268e−19,
                    'link_vel': 1.0129041e−20,
                    'qpos': 1.6398993,
                    'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #74————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.343495637178421,
                                 −0.1483060121536255,
                                 0.48075947165489197],
                    'link_ang_vel': 1.1369268e−19,
                    'link_vel': 1.0129041e−20,
                    'qpos': 1.6398993,
                    'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #75————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.33813759684562683,
                                 −0.15141154825687408,
                                 0.48078298568725586],
                    'link_ang_vel': 1.1369268e−19,
                    'link_vel': 1.0129041e−20,
                    'qpos': 1.6398993,
                    'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #76————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
```

'state_info': {'gripper_pos': [−0.33675527572631836,
                                −0.14471033215522766,
                                0.48430636525154114],
                'link_ang_vel': 1.1369268e−19,
                'link_vel': 1.0129041e−20,
                'qpos': 1.6398993,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #77————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3408561646938324,
                                −0.15277056396007538,
                                0.4901435077190399],
                'link_ang_vel': 1.1369268e−19,
                'link_vel': 1.0129041e−20,
                'qpos': 1.6398993,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #78————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.33671772480010986,
                                −0.15557841956615448,
                                0.4883464574813843],
                'link_ang_vel': 1.1369268e−19,
                'link_vel': 1.0129041e−20,
                'qpos': 1.6398993,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #79————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3350710868835449,
                                −0.1612195372581482,
                                0.4811738133430481],
                'link_ang_vel': 1.1369268e−19,
                'link_vel': 1.0129041e−20,
                'qpos': 1.6398993,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #80————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.33559557795524597,
                                −0.16627994179725647,
                                0.48260343074798584],
                'link_ang_vel': 1.1369268e−19,
                'link_vel': 1.0129041e−20,
                'qpos': 1.6398993,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}

——step #81————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3380168676376343,
                                −0.1681574434041977,
                                0.4832059442996979],
                'link_ang_vel': 1.1369268e−19,
                'link_vel': 1.0129041e−20,
                'qpos': 1.6398993,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #82————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.33881324529647827,
                                −0.15769276022911072,
                                0.4825564920902252],
                'link_ang_vel': 1.1369268e−19,
                'link_vel': 1.0129041e−20,
                'qpos': 1.6398993,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #83————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.34856027364730835,
                                −0.15697415173053741,
                                0.47511056065559387],
                'link_ang_vel': 1.1369268e−19,
                'link_vel': 1.0129041e−20,
                'qpos': 1.6398993,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #84————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3595699965953827,
                                −0.1612350344657898,
                                0.4745956063270569],
                'link_ang_vel': 1.1369268e−19,
                'link_vel': 1.0129041e−20,
                'qpos': 1.6398993,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
——step #85————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3670918941497803,
                                −0.15789267420768738,
                                0.47921669483184814],
                'link_ang_vel': 1.1369268e−19,
                'link_vel': 1.0129041e−20,
                'qpos': 1.6398993,

```
                    'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': False, 'success': False}
----step #86----
{'eval_info': {'object_static': False, 'open_enough': True, 'success': False},
 'state_info': {'gripper_pos': [-0.3640270531177521,
                                -0.16060331463813782,
                                0.47539761662483215],
                'link_ang_vel': 1.1369268e-19,
                'link_vel': 1.0129041e-20,
                'qpos': 1.6398993,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': False, 'open_enough': True, 'success': False}
----step #87----
{'eval_info': {'object_static': True, 'open_enough': True, 'success': True},
 'state_info': {'gripper_pos': [-0.36434754729270935,
                                -0.15645459294319153,
                                0.47727110981941223],
                'link_ang_vel': 1.1369268e-19,
                'link_vel': 1.0129041e-20,
                'qpos': 1.6398993,
                'target_qpos': 1.5607432007789612}}
reward: 5.000000 {'object_static': True, 'open_enough': True, 'success': True}
Bravo! Open the door!
```

*Part 3*

Random shooting is unable to finish the task in 150 timesteps. It is able to partially open the cabinet, but is unable to open it beyond the *target_qpos*. Random shooting thus does not perform as well as CEM-only.

The complete log of random shooting is shown below:

```
——step #0———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3581922948360443,
                                0.015903688967227936,
                                0.5068635940551758],
               'link_ang_vel': 1.4596707e−16,
               'link_vel': 2.7283686e−16,
               'qpos': 3.552714e−18,
               'target_qpos': 1.5607432007789612}}
reward: −0.126908 {'object_static': False, 'open_enough': False, 'success': False}
——step #1———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.34364426136016846,
                                0.02146386168897152,
                                0.4998109042644501],
               'link_ang_vel': 2.1175824e−22,
               'link_vel': 1.5825302e−16,
               'qpos': 0.0,
               'target_qpos': 1.5607432007789612}}
reward: −0.110237 {'object_static': False, 'open_enough': False, 'success': False}
——step #2———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3409403860569,
                                0.023995045572519302,
                                0.4979342222213745],
               'link_ang_vel': 1.4596707e−16,
               'link_vel': 2.7283686e−16,
               'qpos': 3.552714e−18,
               'target_qpos': 1.5607432007789612}}
reward: −0.105150 {'object_static': False, 'open_enough': False, 'success': False}
——step #3———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.36409035325050354,
                                0.024182146415114403,
                                0.4847555160522461],
               'link_ang_vel': 2.1175824e−22,
               'link_vel': 1.5825302e−16,
               'qpos': 0.0,
               'target_qpos': 1.5607432007789612}}
reward: −0.111964 {'object_static': False, 'open_enough': False, 'success': False}
——step #4———
```

{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.36444535851478577,
                                0.024709492921829224,
                                0.46705999970436096],
                'link_ang_vel': 1.4596707e−16,
                'link_vel': 2.7283686e−16,
                'qpos': 3.552714e−18,
                'target_qpos': 1.5607432007789612}}
reward: −0.096360 {'object_static': False, 'open_enough': False, 'success': False}
——step #5———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.36243805289268494,
                                0.043773286044597626,
                                0.44258183240890503],
                'link_ang_vel': 2.1175824e−22,
                'link_vel': 1.5825302e−16,
                'qpos': 0.0,
                'target_qpos': 1.5607432007789612}}
reward: −0.070692 {'object_static': False, 'open_enough': False, 'success': False}
——step #6———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3570637106895447,
                                0.04722841456532478,
                                0.42259034514427185],
                'link_ang_vel': 1.4596707e−16,
                'link_vel': 2.7283686e−16,
                'qpos': 3.552714e−18,
                'target_qpos': 1.5607432007789612}}
reward: −0.067503 {'object_static': False, 'open_enough': False, 'success': False}
——step #7———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.34307199716567993,
                                0.03850547596812248,
                                0.4384555518627167],
                'link_ang_vel': 2.1175824e−22,
                'link_vel': 1.5825302e−16,
                'qpos': 0.0,
                'target_qpos': 1.5607432007789612}}
reward: −0.045164 {'object_static': False, 'open_enough': False, 'success': False}
——step #8———
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.33430734276771545,
                                0.047720301896333694,
                                0.4368930459022522],
                'link_ang_vel': 1.4596707e−16,
                'link_vel': 2.7283686e−16,
                'qpos': 3.552714e−18,
                'target_qpos': 1.5607432007789612}}

45

reward: −0.032866 {'object_static': False, 'open_enough': False, 'success': False}
——step #9————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3221173584461212,
                                0.05757449194788933,
                                0.4145847260951996],
                'link_ang_vel': 6.0817474e−16,
                'link_vel': 2.3166296e−16,
                'qpos': 1.3843095e−18,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #10————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3238162696361542,
                                0.03791205957531929,
                                0.4106377363204956],
                'link_ang_vel': 0.0050471784,
                'link_vel': 0.0004496624,
                'qpos': −4.7226854e−11,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #11————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3246423304080963,
                                0.03418790549039841,
                                0.3913150429725647],
                'link_ang_vel': 0.0046097036,
                'link_vel': 0.00041068703,
                'qpos': −1.9426455e−10,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #12————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.33484551310539246,
                                0.03496858477592468,
                                0.396878719329834],
                'link_ang_vel': 0.013249567,
                'link_vel': 0.0011804282,
                'qpos': −6.87952e−12,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #13————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.33747243881225586,
                                0.02786283940076828,
                                0.4071885049343109],
                'link_ang_vel': 8.855089e−09,
                'link_vel': 7.889163e−10,

```
                        'qpos': 0.0001608853,
                        'target_qpos': 1.5607432007789612}}
reward: 0.439418 {'object_static': False, 'open_enough': False, 'success': False}
——step #14————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.34887969493865967,
                                0.018466083332896233,
                                0.4126488268375397],
                'link_ang_vel': 0.0030207797,
                'link_vel': 0.00026912685,
                'qpos': 4.9112714e−11,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #15————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.35120248794555664,
                                0.022931843996047974,
                                0.4084600806236267],
                'link_ang_vel': 1.0831978e−08,
                'link_vel': 9.650408e−10,
                'qpos': −3.0741945e−12,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #16————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3535732626914978,
                                0.02469903603196144,
                                0.4183937609195709],
                'link_ang_vel': 0.022263031,
                'link_vel': 0.001983455,
                'qpos': −2.747313e−10,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #17————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3490940034389496,
                                0.023014510050415993,
                                0.4037140905857086],
                'link_ang_vel': 0.0022389484,
                'link_vel': 0.00019947204,
                'qpos': −3.3751564e−11,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #18————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.37122514843940735,
                                0.02666468545794487,
                                0.400672048330307],
```

47

                        'link_ang_vel': 6.6071926e−11,
                        'link_vel': 5.8864545e−12,
                        'qpos': 0.0001962574,
                        'target_qpos': 1.5607432007789612}}
reward: 0.439453 {'object_static': False, 'open_enough': False, 'success': False}
——step #19————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.36615481972694397,
                        0.0268126018345356,
                        0.3986026644706726],
                        'link_ang_vel': 0.0063257045,
                        'link_vel': 0.0005635687,
                        'qpos': −4.1582092e−11,
                        'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #20————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3614500164985657,
                        0.028148125857114792,
                        0.3916046917438507],
                        'link_ang_vel': 1.5470323e−06,
                        'link_vel': 1.3782784e−07,
                        'qpos': 3.731842e−11,
                        'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #21————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.36620575189590454,
                        0.031829312443733215,
                        0.3949121832847595],
                        'link_ang_vel': 1.5649156e−09,
                        'link_vel': 1.394211e−10,
                        'qpos': −3.727174e−12,
                        'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #22————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.366100937128067,
                        0.03680988401174545,
                        0.38963332772254944],
                        'link_ang_vel': 0.45341536,
                        'link_vel': 0.040395614,
                        'qpos': 0.0117403995,
                        'target_qpos': 1.5607432007789612}}
reward: 0.450997 {'object_static': True, 'open_enough': False, 'success': False}
——step #23————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.36192864179611206,

```
                                0.04688425734639168,
                                0.38306480646133423],
                    'link_ang_vel': 0.09413489,
                    'link_vel': 0.008386651,
                    'qpos': -4.7869975e-10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #24————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [-0.35842400789260864,
                                0.048655662685632706,
                                0.39076629281044006],
                    'link_ang_vel': 0.6716472,
                    'link_vel': 0.059838343,
                    'qpos': 0.02367192,
                    'target_qpos': 1.5607432007789612}}
reward: 0.462929 {'object_static': False, 'open_enough': False, 'success': False}
——step #25————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [-0.36203858256340027,
                                0.025936637073755264,
                                0.4082179069519043],
                    'link_ang_vel': 0.1678865,
                    'link_vel': 0.014957317,
                    'qpos': 0.0149129955,
                    'target_qpos': 1.5607432007789612}}
reward: 0.454170 {'object_static': False, 'open_enough': False, 'success': False}
——step #26————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [-0.3589830994606018,
                                0.014437868259847164,
                                0.4095565676689148],
                    'link_ang_vel': 2.111806e-09,
                    'link_vel': 1.8814474e-10,
                    'qpos': 0.0149129955,
                    'target_qpos': 1.5607432007789612}}
reward: 0.454170 {'object_static': False, 'open_enough': False, 'success': False}
——step #27————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [-0.3648844361305237,
                                0.01628963276743889,
                                0.4109559953212738],
                    'link_ang_vel': 0.4745048,
                    'link_vel': 0.04227452,
                    'qpos': 0.0021058489,
                    'target_qpos': 1.5607432007789612}}
reward: 0.441363 {'object_static': False, 'open_enough': False, 'success': False}
——step #28————
```

{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3601312041282654,
                               0.025594200938940048,
                               0.39822641015052795],
               'link_ang_vel': 0.0020304746,
               'link_vel': 0.00018089876,
               'qpos': −3.213061e−11,
               'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #29————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.366052508354187,
                               0.014570869505405426,
                               0.383074551820755],
               'link_ang_vel': 8.623906e−09,
               'link_vel': 7.6831996e−10,
               'qpos': 4.9761896e−11,
               'target_qpos': 1.5607432007789612}}
reward: −0.029604 {'object_static': False, 'open_enough': False, 'success': False}
——step #30————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3636731207370758,
                               0.018744144588708878,
                               0.3760366141796112],
               'link_ang_vel': 0.008466706,
               'link_vel': 0.00075431453,
               'qpos': −1.2151298e−10,
               'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #31————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.36847779154777527,
                               0.030750872567296028,
                               0.3770454227924347],
               'link_ang_vel': 0.013607858,
               'link_vel': 0.0012123492,
               'qpos': 3.308659e−11,
               'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #32————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3607128858566284,
                               0.013469159603118896,
                               0.37726348638534546],
               'link_ang_vel': 1.3304735e−08,
               'link_vel': 1.1853436e−09,
               'qpos': 8.008404e−11,
               'target_qpos': 1.5607432007789612}}

reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #33————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3668595850467682,
                                −0.004057401791214943,
                                0.38899824023246765],
                'link_ang_vel': 0.010532956,
                'link_vel': 0.00093840045,
                'qpos': −3.776003e−11,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #34————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3710354268550873,
                                −0.013886361382901669,
                                0.39100852608680725],
                'link_ang_vel': 0.23514678,
                'link_vel': 0.02094966,
                'qpos': 0.012652888,
                'target_qpos': 1.5607432007789612}}
reward: 0.451910 {'object_static': False, 'open_enough': False, 'success': False}
——step #35————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3725540041923523,
                                −0.023434946313500404,
                                0.38987600803375244],
                'link_ang_vel': 2.108191e−09,
                'link_vel': 1.878227e−10,
                'qpos': 0.011219773,
                'target_qpos': 1.5607432007789612}}
reward: 0.450477 {'object_static': False, 'open_enough': False, 'success': False}
——step #36————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3785015344619751,
                                −0.02824302203953266,
                                0.3796512484550476],
                'link_ang_vel': 2.8133265e−09,
                'link_vel': 2.5064442e−10,
                'qpos': 0.011219773,
                'target_qpos': 1.5607432007789612}}
reward: 0.450477 {'object_static': False, 'open_enough': False, 'success': False}
——step #37————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.38762006163597107,
                                −0.021070171147584915,
                                0.3768449127674103],
                'link_ang_vel': 3.069783e−19,
                'link_vel': 2.7349266e−20,

```
                   'qpos': 0.011219773,
                   'target_qpos': 1.5607432007789612}}
reward: 0.450477 {'object_static': False, 'open_enough': False, 'success': False}
——step #38————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.38191667199134827,
                                −0.033741988241672516,
                                0.3834191858768463],
                'link_ang_vel': 0.18116751,
                'link_vel': 0.016140547,
                'qpos': 0.004309062,
                'target_qpos': 1.5607432007789612}}
reward: 0.443566 {'object_static': False, 'open_enough': False, 'success': False}
——step #39————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3906019628047943,
                                −0.034454893320798874,
                                0.38316115736961365],
                'link_ang_vel': 3.159087e−06,
                'link_vel': 2.8167082e−07,
                'qpos': 6.193296e−11,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #40————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.38474440574645996,
                                −0.04053419828414917,
                                0.3868696987628937],
                'link_ang_vel': 0.015069615,
                'link_vel': 0.00134258,
                'qpos': −3.5849514e−11,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #41————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.37912681698799133,
                                −0.03887270390987396,
                                0.37989750504493713],
                'link_ang_vel': 0.017099177,
                'link_vel': 0.0015233972,
                'qpos': −4.888803e−11,
                'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #42————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.37680912017822266,
                                −0.033343032002449036,
                                0.36654922366142273],
```

                    'link_ang_vel': 0.005296012,
                    'link_vel': 0.00047183156,
                    'qpos': 1.4328638e−10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #43————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3785361349582672,
                                −0.03912985324859619,
                                0.35938122868537903],
                    'link_ang_vel': 0.04337053,
                    'link_vel': 0.0038639607,
                    'qpos': −2.134102e−10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': False, 'open_enough': False, 'success': False}
——step #44————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.39242854714393616,
                                −0.0507485531270504,
                                0.35796260833740234],
                    'link_ang_vel': 0.012838714,
                    'link_vel': 0.0011438248,
                    'qpos': −7.371183e−11,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #45————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3964303433895111,
                                −0.053612519055604935,
                                0.34549713134765625],
                    'link_ang_vel': 0.01133714,
                    'link_vel': 0.0010100467,
                    'qpos': −4.2046838e−11,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #46————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.40323859453201294,
                                −0.05310425907373428,
                                0.34473365016937256],
                    'link_ang_vel': 0.009330463,
                    'link_vel': 0.0008312681,
                    'qpos': −2.714567e−11,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #47————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.38950902223587036,

                                    −0.03779567405581474,
                                    0.33288198709487915],
                    'link_ang_vel': 0.009029961,
                    'link_vel': 0.000804496,
                    'qpos': −4.5772133e−11,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #48————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.38343867659568787,
                                    −0.03638727217912674,
                                    0.3258908689022064],
                    'link_ang_vel': 0.014763494,
                    'link_vel': 0.0013153069,
                    'qpos': 3.0596303e−11,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #49————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3914368152618408,
                                    −0.03445115312933922,
                                    0.31672367453575134],
                    'link_ang_vel': 1.8673822e−09,
                    'link_vel': 1.6636867e−10,
                    'qpos': 3.7422482e−10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #50————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3812992572784424,
                                    −0.034499965608119965,
                                    0.33702462911605835],
                    'link_ang_vel': 1.057237e−05,
                    'link_vel': 9.41912e−07,
                    'qpos': 1.0852878e−10,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #51————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3912683129310608,
                                    −0.028778526932001114,
                                    0.34086355566978455],
                    'link_ang_vel': 3.498154e−10,
                    'link_vel': 3.1165927e−11,
                    'qpos': 8.7200296e−11,
                    'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #52————

{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.39044445753097534,
                                 −0.02176758088171482,
                                 0.3332774341106415],
                 'link_ang_vel': 0.01402724,
                 'link_vel': 0.0012497128,
                 'qpos': −8.0394656e−11,
                 'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #53————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3841401934623718,
                                 −0.011256481520831585,
                                 0.343476265668869],
                 'link_ang_vel': 1.7033275e−05,
                 'link_vel': 1.517533e−06,
                 'qpos': −2.0239783e−10,
                 'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #54————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3858678340911865,
                                 −0.010893675498664379,
                                 0.34008920192718506],
                 'link_ang_vel': 7.342976e−10,
                 'link_vel': 6.542025e−11,
                 'qpos': 7.855997e−12,
                 'target_qpos': 1.5607432007789612}}
reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #55————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.39262261986732483,
                                 −0.016443662345409393,
                                 0.35029202699661255],
                 'link_ang_vel': 3.8505085e−08,
                 'link_vel': 3.4304886e−09,
                 'qpos': 0.000997926,
                 'target_qpos': 1.5607432007789612}}
reward: 0.440255 {'object_static': True, 'open_enough': False, 'success': False}
——step #56————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3875933289527893,
                                 −0.022672312334179878,
                                 0.35226374864578247],
                 'link_ang_vel': 3.574374e−11,
                 'link_vel': 3.1844952e−12,
                 'qpos': 3.1840969e−12,
                 'target_qpos': 1.5607432007789612}}

55

reward: 0.439257 {'object_static': True, 'open_enough': False, 'success': False}
——step #57————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.39544761180877686,
                                 0.00035303860204294324,
                                 0.35617169737815857],
                'link_ang_vel': 0.43585798,
                'link_vel': 0.03883139,
                'qpos': 0.0054449337,
                'target_qpos': 1.5607432007789612}}
reward: 0.444702 {'object_static': True, 'open_enough': False, 'success': False}
——step #58————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3907789885997772,
                                 0.007082258351147175,
                                 0.3510689437389374],
                'link_ang_vel': 1.521232e−10,
                'link_vel': 1.3552942e−11,
                'qpos': 0.006670666,
                'target_qpos': 1.5607432007789612}}
reward: 0.445928 {'object_static': True, 'open_enough': False, 'success': False}
——step #59————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.39808201789855957,
                                 0.017971739172935486,
                                 0.3511478006839752],
                'link_ang_vel': 0.053840965,
                'link_vel': 0.0047967895,
                'qpos': 0.0072528105,
                'target_qpos': 1.5607432007789612}}
reward: 0.446510 {'object_static': True, 'open_enough': False, 'success': False}
——step #60————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.401725172996521,
                                 0.02123379148542881,
                                 0.34963464736938477],
                'link_ang_vel': 1.534889e−19,
                'link_vel': 1.3674607e−20,
                'qpos': 0.0072528105,
                'target_qpos': 1.5607432007789612}}
reward: 0.446510 {'object_static': True, 'open_enough': False, 'success': False}
——step #61————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.3902912437915802,
                                 0.01250887755304575,
                                 0.3581339418888092],
                'link_ang_vel': 0.2531514,
                'link_vel': 0.022553721,

                    'qpos': 0.0025927036,
                    'target_qpos': 1.5607432007789612}}
reward: 0.441849 {'object_static': True, 'open_enough': False, 'success': False}
——step #62————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3925611078739166,
                                0.00583953270688653,
                                0.3573235273361206],
                'link_ang_vel': 0.26249182,
                'link_vel': 0.023385882,
                'qpos': 0.014988233,
                'target_qpos': 1.5607432007789612}}
reward: 0.454245 {'object_static': True, 'open_enough': False, 'success': False}
——step #63————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3898742198944092,
                                0.0030821876134723425,
                                0.333911657333374],
                'link_ang_vel': 4.1885265e−19,
                'link_vel': 2.220073e−16,
                'qpos': 0.015047942,
                'target_qpos': 1.5607432007789612}}
reward: 0.454305 {'object_static': True, 'open_enough': False, 'success': False}
——step #64————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.38837358355522156,
                                0.004921527113765478,
                                0.34714823961257935],
                'link_ang_vel': 3.9816664e−11,
                'link_vel': 3.5476496e−12,
                'qpos': 0.015047942,
                'target_qpos': 1.5607432007789612}}
reward: 0.454305 {'object_static': True, 'open_enough': False, 'success': False}
——step #65————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.39919713139533997,
                                −0.005613610148429871,
                                0.35604923963546753],
                'link_ang_vel': 0.3895766,
                'link_vel': 0.03470811,
                'qpos': 0.020229436,
                'target_qpos': 1.5607432007789612}}
reward: 0.459486 {'object_static': True, 'open_enough': False, 'success': False}
——step #66————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.39721041917800903,
                                −0.01168528851121664,
                                0.3578110933303833],

```
                    'link_ang_vel': 0.019497845,
                    'link_vel': 0.0017370987,
                    'qpos': 0.022231912,
                    'target_qpos': 1.5607432007789612}}
reward: 0.461489 {'object_static': True, 'open_enough': False, 'success': False}
——step #67————
{'eval_info': {'object_static': True, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.39654576778411865,
                    −0.0135938236489892,
                    0.36044952273368835],
                    'link_ang_vel': 0.14624617,
                    'link_vel': 0.01302934,
                    'qpos': 0.027522456,
                    'target_qpos': 1.5607432007789612}}
reward: 0.466779 {'object_static': True, 'open_enough': False, 'success': False}
——step #68————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4019901752471924,
                    −0.013244770467281342,
                    0.3598692715167999],
                    'link_ang_vel': 1.0391803,
                    'link_vel': 0.09258268,
                    'qpos': 0.06710821,
                    'target_qpos': 1.5607432007789612}}
reward: 0.506365 {'object_static': False, 'open_enough': False, 'success': False}
——step #69————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3911713659763336,
                    −0.014872504398226738,
                    0.3478974997997284],
                    'link_ang_vel': 0.7636184,
                    'link_vel': 0.06803222,
                    'qpos': 0.063303314,
                    'target_qpos': 1.5607432007789612}}
reward: 0.502560 {'object_static': False, 'open_enough': False, 'success': False}
——step #70————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.40061619877815247,
                    −0.014866883866488934,
                    0.35955461859703064],
                    'link_ang_vel': 1.2136209,
                    'link_vel': 0.10812408,
                    'qpos': 0.0587259,
                    'target_qpos': 1.5607432007789612}}
reward: 0.497983 {'object_static': False, 'open_enough': False, 'success': False}
——step #71————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.40155869722366333,
```

$$-0.016635365784168243,$$
$$0.35440123081207275],$$
'link_ang_vel': 7.299243e−17,
'link_vel': 8.563247e−17,
'qpos': 0.048456945,
'target_qpos': 1.5607432007789612}}
reward: 0.487714 {'object_static': False, 'open_enough': False, 'success': False}
——step #72————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
'state_info': {'gripper_pos': [−0.40603306889533997,
                               −0.010018623434007168,
                               0.36410900950431824],
              'link_ang_vel': 0.2298751,
              'link_vel': 0.020479996,
              'qpos': 0.05442828,
              'target_qpos': 1.5607432007789612}}
reward: 0.493685 {'object_static': False, 'open_enough': False, 'success': False}
——step #73————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
'state_info': {'gripper_pos': [−0.3903183043003082,
                               0.004364179912954569,
                               0.36035141348838806],
              'link_ang_vel': 1.0104408e−08,
              'link_vel': 9.002202e−10,
              'qpos': 0.051171098,
              'target_qpos': 1.5607432007789612}}
reward: 0.490428 {'object_static': False, 'open_enough': False, 'success': False}
——step #74————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
'state_info': {'gripper_pos': [−0.40259218215942383,
                               0.00517776794731617,
                               0.3589804470539093],
              'link_ang_vel': 1.0989211,
              'link_vel': 0.09790513,
              'qpos': 0.09607208,
              'target_qpos': 1.5607432007789612}}
reward: 0.535329 {'object_static': False, 'open_enough': False, 'success': False}
——step #75————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
'state_info': {'gripper_pos': [−0.4111420810222626,
                               0.012809344567358494,
                               0.3581973612308502],
              'link_ang_vel': 0.006095512,
              'link_vel': 0.0005430604,
              'qpos': 0.09877377,
              'target_qpos': 1.5607432007789612}}
reward: 0.538031 {'object_static': False, 'open_enough': False, 'success': False}
——step #76————

{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.3955903947353363,
                                0.009746624156832695,
                                0.3454529345035553],
                'link_ang_vel': 0.47336102,
                'link_vel': 0.04217264,
                'qpos': 0.10255289,
                'target_qpos': 1.5607432007789612}}
reward: 0.541810 {'object_static': False, 'open_enough': False, 'success': False}
——step #77————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.403914213180542,
                                0.012482210993766785,
                                0.33855998516082764],
                'link_ang_vel': 0.063458934,
                'link_vel': 0.005653673,
                'qpos': 0.12756063,
                'target_qpos': 1.5607432007789612}}
reward: 0.566817 {'object_static': False, 'open_enough': False, 'success': False}
——step #78————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.39908167719841003,
                                −0.00032058777287602425,
                                0.34471043944358826],
                'link_ang_vel': 0.46988434,
                'link_vel': 0.041862875,
                'qpos': 0.117031164,
                'target_qpos': 1.5607432007789612}}
reward: 0.556288 {'object_static': False, 'open_enough': False, 'success': False}
——step #79————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.40254610776901245,
                                0.003918416798114777,
                                0.3417372405529022],
                'link_ang_vel': 0.11893807,
                'link_vel': 0.010596415,
                'qpos': 0.12135107,
                'target_qpos': 1.5607432007789612}}
reward: 0.560608 {'object_static': False, 'open_enough': False, 'success': False}
——step #80————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.40880247950553894,
                                −0.004213742446154356,
                                0.34603092074394226],
                'link_ang_vel': 0.7163212,
                'link_vel': 0.063818425,
                'qpos': 0.15225895,
                'target_qpos': 1.5607432007789612}}

reward: 0.591516 {'object_static': False, 'open_enough': False, 'success': False}
——step *#81*————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.41117775440216064,
                                −0.012819216586649418,
                                0.3469604551792145],
                'link_ang_vel': 0.84038424,
                'link_vel': 0.0748715,
                'qpos': 0.1938512,
                'target_qpos': 1.5607432007789612}}
reward: 0.633108 {'object_static': False, 'open_enough': False, 'success': False}
——step *#82*————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.40804043412208557,
                                −0.0032655808608978987,
                                0.34600013494491577],
                'link_ang_vel': 1.4884372e−10,
                'link_vel': 1.3260758e−11,
                'qpos': 0.19678044,
                'target_qpos': 1.5607432007789612}}
reward: 0.636037 {'object_static': False, 'open_enough': False, 'success': False}
——step *#83*————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4084296226501465,
                                0.003913560416549444,
                                0.3376409113407135],
                'link_ang_vel': 0.0,
                'link_vel': 0.0,
                'qpos': 0.19678044,
                'target_qpos': 1.5607432007789612}}
reward: 0.636037 {'object_static': False, 'open_enough': False, 'success': False}
——step *#84*————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.41359931230545044,
                                0.011515077203512192,
                                0.3489120602607727],
                'link_ang_vel': 0.40258667,
                'link_vel': 0.035867203,
                'qpos': 0.20340839,
                'target_qpos': 1.5607432007789612}}
reward: 0.642665 {'object_static': False, 'open_enough': False, 'success': False}
——step *#85*————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4162721633911133,
                                0.005780125968158245,
                                0.35130029916763306],
                'link_ang_vel': 0.6256197,
                'link_vel': 0.055737656,

61

```
                    'qpos': 0.23458056,
                    'target_qpos': 1.5607432007789612}}
reward: 0.673837 {'object_static': False, 'open_enough': False, 'success': False}
——step #86————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.40983203053474426,
                                 −0.00023772148415446281,
                                 0.34924623370170593],
                    'link_ang_vel': 2.3798614e−19,
                    'link_vel': 2.1202707e−20,
                    'qpos': 0.23753166,
                    'target_qpos': 1.5607432007789612}}
reward: 0.676788 {'object_static': False, 'open_enough': False, 'success': False}
——step #87————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.40396547317504883,
                                 −0.003059553913772106,
                                 0.346608430147171],
                    'link_ang_vel': 0.9090954,
                    'link_vel': 0.08099312,
                    'qpos': 0.1995301,
                    'target_qpos': 1.5607432007789612}}
reward: 0.638787 {'object_static': False, 'open_enough': False, 'success': False}
——step #88————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4046420156955719,
                                 −0.015325170941650867,
                                 0.3565139174461365],
                    'link_ang_vel': 0.75375515,
                    'link_vel': 0.06715346,
                    'qpos': 0.21557036,
                    'target_qpos': 1.5607432007789612}}
reward: 0.654827 {'object_static': False, 'open_enough': False, 'success': False}
——step #89————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4081641137599945,
                                 −0.006333020515739918,
                                 0.3553423583507538],
                    'link_ang_vel': 0.37227407,
                    'link_vel': 0.03316659,
                    'qpos': 0.2347166,
                    'target_qpos': 1.5607432007789612}}
reward: 0.673973 {'object_static': False, 'open_enough': False, 'success': False}
——step #90————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4065789580345154,
                                 −0.0031229460146278143,
                                 0.3541222810745239],
```

                    'link_ang_vel': 4.8045067e−08,
                    'link_vel': 4.2804236e−09,
                    'qpos': 0.26077688,
                    'target_qpos': 1.5607432007789612}}
reward: 0.700034 {'object_static': False, 'open_enough': False, 'success': False}
——step #91————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.40501007437705994,
                                −0.018772050738334656,
                                0.346314013004303],
                    'link_ang_vel': 7.299338e−17,
                    'link_vel': 8.318211e−17,
                    'qpos': 0.260623,
                    'target_qpos': 1.5607432007789612}}
reward: 0.699880 {'object_static': False, 'open_enough': False, 'success': False}
——step #92————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.41654425859451294,
                                −0.0061183650977909565,
                                0.35438162088394165],
                    'link_ang_vel': 0.7949839,
                    'link_vel': 0.07082665,
                    'qpos': 0.2796191,
                    'target_qpos': 1.5607432007789612}}
reward: 0.718876 {'object_static': False, 'open_enough': False, 'success': False}
——step #93————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4166066646575928,
                                −0.002208384219557047,
                                0.3500266671180725],
                    'link_ang_vel': 1.3978145e−09,
                    'link_vel': 1.2453388e−10,
                    'qpos': 0.28287178,
                    'target_qpos': 1.5607432007789612}}
reward: 0.722129 {'object_static': False, 'open_enough': False, 'success': False}
——step #94————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.42783910036087036,
                                0.007849023677408695,
                                0.34807708859443665],
                    'link_ang_vel': 0.0,
                    'link_vel': 0.0,
                    'qpos': 0.28287178,
                    'target_qpos': 1.5607432007789612}}
reward: 0.722129 {'object_static': False, 'open_enough': False, 'success': False}
——step #95————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.43304386734962463,

                                    −0.0028827909845858812,
                                    0.335928738117218],
                    'link_ang_vel': 1.895181e−10,
                    'link_vel': 1.6884437e−11,
                    'qpos': 0.28289527,
                    'target_qpos': 1.5607432007789612}}
reward: 0.722152 {'object_static': False, 'open_enough': False, 'success': False}
——step #96————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4101454019546509,
                                    −0.02057664841413498,
                                    0.355654239654541],
                    'link_ang_vel': 2.6117764e−08,
                    'link_vel': 2.3268796e−09,
                    'qpos': 0.2828245,
                    'target_qpos': 1.5607432007789612}}
reward: 0.722081 {'object_static': False, 'open_enough': False, 'success': False}
——step #97————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4062291383743286,
                                    −0.036149680614471436,
                                    0.35649824142456055],
                    'link_ang_vel': 0.39854574,
                    'link_vel': 0.035507187,
                    'qpos': 0.30300385,
                    'target_qpos': 1.5607432007789612}}
reward: 0.742261 {'object_static': False, 'open_enough': False, 'success': False}
——step #98————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.40709927678108215,
                                    −0.03443855792284012,
                                    0.3595235049724579],
                    'link_ang_vel': 0.34184712,
                    'link_vel': 0.0304558,
                    'qpos': 0.31997445,
                    'target_qpos': 1.5607432007789612}}
reward: 0.759231 {'object_static': False, 'open_enough': False, 'success': False}
——step #99————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4186147153377533,
                                    −0.01671251654624939,
                                    0.35039010643959045],
                    'link_ang_vel': 0.018330624,
                    'link_vel': 0.0016331092,
                    'qpos': 0.32526636,
                    'target_qpos': 1.5607432007789612}}
reward: 0.764523 {'object_static': False, 'open_enough': False, 'success': False}
——step #100————

{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4194265305995941,
                                −0.0010024785296991467,
                                0.3550565242767334],
                'link_ang_vel': 0.24513507,
                'link_vel': 0.02183955,
                'qpos': 0.3357423,
                'target_qpos': 1.5607432007789612}}
reward: 0.774999 {'object_static': False, 'open_enough': False, 'success': False}
——step #101————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4209015667438507,
                                0.004658353514969349,
                                0.3519762456417084],
                'link_ang_vel': 0.7455721,
                'link_vel': 0.06642447,
                'qpos': 0.36984655,
                'target_qpos': 1.5607432007789612}}
reward: 0.809103 {'object_static': False, 'open_enough': False, 'success': False}
——step #102————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.42465534806251526,
                                0.013750172220170498,
                                0.34967517852783203],
                'link_ang_vel': 2.2116623e−08,
                'link_vel': 1.97041e−09,
                'qpos': 0.37464792,
                'target_qpos': 1.5607432007789612}}
reward: 0.813905 {'object_static': False, 'open_enough': False, 'success': False}
——step #103————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.42215898633003235,
                                0.003621109062805772,
                                0.3311975598335266],
                'link_ang_vel': 1.5865004e−10,
                'link_vel': 1.4134664e−11,
                'qpos': 0.37478793,
                'target_qpos': 1.5607432007789612}}
reward: 0.814045 {'object_static': False, 'open_enough': False, 'success': False}
——step #104————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.41953760385513306,
                                0.0003011227527167648,
                                0.3450917601585388],
                'link_ang_vel': 4.8603747e−19,
                'link_vel': 2.2200417e−16,
                'qpos': 0.37478793,
                'target_qpos': 1.5607432007789612}}

reward: 0.814045 {'object_static': False, 'open_enough': False, 'success': False}
——step *#105*————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.42774468660354614,
                                0.01736706681549549,
                                0.3547089099884033],
                'link_ang_vel': 0.14757709,
                'link_vel': 0.013147915,
                'qpos': 0.3757861,
                'target_qpos': 1.5607432007789612}}
reward: 0.815043 {'object_static': False, 'open_enough': False, 'success': False}
——step *#106*————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4273635149002075,
                                0.022420991212129593,
                                0.355012178421020],
                'link_ang_vel': 0.015433017,
                'link_vel': 0.0013749556,
                'qpos': 0.37792733,
                'target_qpos': 1.5607432007789612}}
reward: 0.817184 {'object_static': False, 'open_enough': False, 'success': False}
——step *#107*————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4388783574104309,
                                0.0027523734606802464,
                                0.35987746715545654],
                'link_ang_vel': 0.77091545,
                'link_vel': 0.06868236,
                'qpos': 0.41056326,
                'target_qpos': 1.5607432007789612}}
reward: 0.849820 {'object_static': False, 'open_enough': False, 'success': False}
——step *#108*————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4417724311351776,
                                −0.007872454822063446,
                                0.36018502712249756],
                'link_ang_vel': 0.34476453,
                'link_vel': 0.030715711,
                'qpos': 0.43199694,
                'target_qpos': 1.5607432007789612}}
reward: 0.871254 {'object_static': False, 'open_enough': False, 'success': False}
——step *#109*————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4432264566421509,
                                −0.01746596395969391,
                                0.3594743311405182],
                'link_ang_vel': 2.5680165e−09,
                'link_vel': 2.2878927e−10,

```
                    'qpos': 0.44054964,
                    'target_qpos': 1.5607432007789612}}
reward: 0.879806 {'object_static': False, 'open_enough': False, 'success': False}
——step #110————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4375166893005371,
                                −0.022249359637498856,
                                0.3504030704498291],
                  'link_ang_vel': 1.03909374e−10,
                  'link_vel': 9.257397e−12,
                  'qpos': 0.44054964,
                  'target_qpos': 1.5607432007789612}}
reward: 0.879806 {'object_static': False, 'open_enough': False, 'success': False}
——step #111————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4289212226867676,
                                −0.012606448493897915,
                                0.34195250272750854],
                  'link_ang_vel': 4.860383e−19,
                  'link_vel': 2.482805e−16,
                  'qpos': 0.44054964,
                  'target_qpos': 1.5607432007789612}}
reward: 0.879806 {'object_static': False, 'open_enough': False, 'success': False}
——step #112————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.42049023509025574,
                                −0.007464323192834854,
                                0.3399452567100525],
                  'link_ang_vel': 2.406822e−09,
                  'link_vel': 2.1442827e−10,
                  'qpos': 0.44054964,
                  'target_qpos': 1.5607432007789612}}
reward: 0.879806 {'object_static': False, 'open_enough': False, 'success': False}
——step #113————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4277857542037964,
                                0.004196478053927422,
                                0.3572733998298645],
                  'link_ang_vel': 0.00096922,
                  'link_vel': 8.634959e−05,
                  'qpos': 0.43676358,
                  'target_qpos': 1.5607432007789612}}
reward: 0.876020 {'object_static': False, 'open_enough': False, 'success': False}
——step #114————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.43508678674697876,
                                −0.0014088499592617154,
                                0.35984256863594055],
```

67

```
                          'link_ang_vel': 0.86819005,
                          'link_vel': 0.07734876,
                          'qpos': 0.47923526,
                          'target_qpos': 1.5607432007789612}}
reward: 0.918492 {'object_static': False, 'open_enough': False, 'success': False}
——step #115————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.43473896384239197,
                                -0.004884742666035891,
                                0.35841795802116394],
                'link_ang_vel': 9.377307e-09,
                'link_vel': 8.3544127e-10,
                'qpos': 0.48309556,
                'target_qpos': 1.5607432007789612}}
reward: 0.922352 {'object_static': False, 'open_enough': False, 'success': False}
——step #116————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.4233589768409729,
                                -0.022334493696689606,
                                0.3626589775085449],
                'link_ang_vel': 0.39874455,
                'link_vel': 0.035524897,
                'qpos': 0.4987364,
                'target_qpos': 1.5607432007789612}}
reward: 0.937993 {'object_static': False, 'open_enough': False, 'success': False}
——step #117————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.42678508162498474,
                                -0.0215969979763031,
                                0.3640475273132324],
                'link_ang_vel': 0.12005636,
                'link_vel': 0.010696041,
                'qpos': 0.50695485,
                'target_qpos': 1.5607432007789612}}
reward: 0.946212 {'object_static': False, 'open_enough': False, 'success': False}
——step #118————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.43146756291389465,
                                -0.024146510288119316,
                                0.35713109374046326],
                'link_ang_vel': 1.647118e-10,
                'link_vel': 1.4674484e-11,
                'qpos': 0.5077637,
                'target_qpos': 1.5607432007789612}}
reward: 0.947021 {'object_static': False, 'open_enough': False, 'success': False}
——step #119————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.42238855361938477,
```

                                    −0.003147088922560215,
                                    0.3533412218093872],
                    'link_ang_vel': 0.9677992,
                    'link_vel': 0.08622312,
                    'qpos': 0.49544516,
                    'target_qpos': 1.5607432007789612}}
reward: 0.934702 {'object_static': False, 'open_enough': False, 'success': False}
——step #120————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4314095973968506,
                                    −0.01798136718571186,
                                    0.3620460033416748],
                    'link_ang_vel': 0.927807,
                    'link_vel': 0.08266016,
                    'qpos': 0.5163025,
                    'target_qpos': 1.5607432007789612}}
reward: 0.955559 {'object_static': False, 'open_enough': False, 'success': False}
——step #121————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4342472553253174,
                                    −0.02574227750301361,
                                    0.35422301292419434],
                    'link_ang_vel': 3.960477e−19,
                    'link_vel': 3.5284724e−20,
                    'qpos': 0.5215778,
                    'target_qpos': 1.5607432007789612}}
reward: 0.960835 {'object_static': False, 'open_enough': False, 'success': False}
——step #122————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4363066256046295,
                                    −0.013662189245223999,
                                    0.3574863076210022],
                    'link_ang_vel': 7.0743233e−09,
                    'link_vel': 6.302644e−10,
                    'qpos': 0.5215778,
                    'target_qpos': 1.5607432007789612}}
reward: 0.960835 {'object_static': False, 'open_enough': False, 'success': False}
——step #123————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
  'state_info': {'gripper_pos': [−0.4347410202026367,
                                    −0.024429993703961372,
                                    0.3481115996837616],
                    'link_ang_vel': 5.215197e−10,
                    'link_vel': 4.646305e−11,
                    'qpos': 0.52180636,
                    'target_qpos': 1.5607432007789612}}
reward: 0.961063 {'object_static': False, 'open_enough': False, 'success': False}
——step #124————

{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4273887276649475,
                                −0.021222418174147606,
                                0.35476770997047424],
                'link_ang_vel': 1.1223465,
                'link_vel': 0.09999214,
                'qpos': 0.546684,
                'target_qpos': 1.5607432007789612}}
reward: 0.985941 {'object_static': False, 'open_enough': False, 'success': False}
——step #125————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4340195655822754,
                                −0.0039062430150806904,
                                0.35181763768196106],
                'link_ang_vel': 2.5985867e−09,
                'link_vel': 2.3151292e−10,
                'qpos': 0.57696253,
                'target_qpos': 1.5607432007789612}}
reward: 1.016219 {'object_static': False, 'open_enough': False, 'success': False}
——step #126————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4477367401123047,
                                −0.009263820014894009,
                                0.35080593824386597],
                'link_ang_vel': 6.445926e−10,
                'link_vel': 5.742803e−11,
                'qpos': 0.5772839,
                'target_qpos': 1.5607432007789612}}
reward: 1.016541 {'object_static': False, 'open_enough': False, 'success': False}
——step #127————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.43434447050094604,
                                −0.0007133990293368697,
                                0.35250598192214966],
                'link_ang_vel': 8.50614e−11,
                'link_vel': 7.578209e−12,
                'qpos': 0.5772839,
                'target_qpos': 1.5607432007789612}}
reward: 1.016541 {'object_static': False, 'open_enough': False, 'success': False}
——step #128————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4388779401779175,
                                −0.00041729118674993515,
                                0.3614138662815094],
                'link_ang_vel': 1.4043057e−09,
                'link_vel': 1.2511205e−10,
                'qpos': 0.5772839,
                'target_qpos': 1.5607432007789612}}

```
reward: 1.016541 {'object_static': False, 'open_enough': False, 'success': False}
----step #129----
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.44324791431427,
                                -0.009250102564692497,
                                0.36506205797195435],
                'link_ang_vel': 0.54404503,
                'link_vel': 0.048470043,
                'qpos': 0.6018244,
                'target_qpos': 1.5607432007789612}}
reward: 1.041081 {'object_static': False, 'open_enough': False, 'success': False}
----step #130----
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.4377534091472626,
                                0.004256746731698513,
                                0.36506837606430054],
                'link_ang_vel': 0.68263406,
                'link_vel': 0.06081718,
                'qpos': 0.59038126,
                'target_qpos': 1.5607432007789612}}
reward: 1.029638 {'object_static': False, 'open_enough': False, 'success': False}
----step #131----
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.44268766045570374,
                                0.00038323955959640443,
                                0.3631941080093384],
                'link_ang_vel': 7.055843e-09,
                'link_vel': 6.2861827e-10,
                'qpos': 0.59068155,
                'target_qpos': 1.5607432007789612}}
reward: 1.029938 {'object_static': False, 'open_enough': False, 'success': False}
----step #132----
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.44658324122428894,
                                -0.016940496861934662,
                                0.3598318099975586],
                'link_ang_vel': 0.6641619,
                'link_vel': 0.059171434,
                'qpos': 0.6159891,
                'target_qpos': 1.5607432007789612}}
reward: 1.055246 {'object_static': False, 'open_enough': False, 'success': False}
----step #133----
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [-0.4522811472415924,
                                -0.038531459867954254,
                                0.3475419282913208],
                'link_ang_vel': 1.2152447,
                'link_vel': 0.108268626,
```

```
                    'qpos': 0.65030164,
                    'target_qpos': 1.5607432007789612}}
reward: 1.089558 {'object_static': False, 'open_enough': False, 'success': False}
——step #134————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.45118510723114014,
                                −0.04421181231737137,
                                0.3330881893634796],
                'link_ang_vel': 0.116166495,
                'link_vel': 0.010349485,
                'qpos': 0.6511741,
                'target_qpos': 1.5607432007789612}}
reward: 1.090431 {'object_static': False, 'open_enough': False, 'success': False}
——step #135————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4491036832332611,
                                −0.049805160611867905,
                                0.32602623105049133],
                'link_ang_vel': 0.50213224,
                'link_vel': 0.044735905,
                'qpos': 0.6294221,
                'target_qpos': 1.5607432007789612}}
reward: 1.068679 {'object_static': False, 'open_enough': False, 'success': False}
——step #136————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.45237937569618225,
                                −0.0582159161567688,
                                0.31513988971710205],
                'link_ang_vel': 0.26077586,
                'link_vel': 0.023233008,
                'qpos': 0.63389605,
                'target_qpos': 1.5607432007789612}}
reward: 1.073153 {'object_static': False, 'open_enough': False, 'success': False}
——step #137————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.43685808777809143,
                                −0.05496865510940552,
                                0.32436686754226685],
                'link_ang_vel': 4.675108e−10,
                'link_vel': 4.165139e−11,
                'qpos': 0.6340652,
                'target_qpos': 1.5607432007789612}}
reward: 1.073322 {'object_static': False, 'open_enough': False, 'success': False}
——step #138————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.45365777611732483,
                                −0.034040313214063644,
                                0.31329938769340515],
```

                    'link_ang_vel': 2.2802094e−11,
                    'link_vel': 2.0314746e−12,
                    'qpos': 0.63395995,
                    'target_qpos': 1.5607432007789612}}
reward: 1.073217 {'object_static': False, 'open_enough': False, 'success': False}
——step #139————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4522915780544281,
                                −0.024857264012098312,
                                0.3032587766647339],
                    'link_ang_vel': 0.35948652,
                    'link_vel': 0.03202732,
                    'qpos': 0.6304917,
                    'target_qpos': 1.5607432007789612}}
reward: 1.069748 {'object_static': False, 'open_enough': False, 'success': False}
——step #140————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4530022144317627,
                                −0.03235486149787903,
                                0.3210184574127197],
                    'link_ang_vel': 0.0,
                    'link_vel': 0.0,
                    'qpos': 0.6274533,
                    'target_qpos': 1.5607432007789612}}
reward: 1.066710 {'object_static': False, 'open_enough': False, 'success': False}
——step #141————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4549715518951416,
                                −0.031086228787899017,
                                0.33377596735954285],
                    'link_ang_vel': 0.108841315,
                    'link_vel': 0.009696873,
                    'qpos': 0.6286767,
                    'target_qpos': 1.5607432007789612}}
reward: 1.067934 {'object_static': False, 'open_enough': False, 'success': False}
——step #142————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.45184510946273804,
                                −0.04047059640288353,
                                0.32808610796928406],
                    'link_ang_vel': 0.64429164,
                    'link_vel': 0.057401173,
                    'qpos': 0.6494723,
                    'target_qpos': 1.5607432007789612}}
reward: 1.088729 {'object_static': False, 'open_enough': False, 'success': False}
——step #143————
{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.44970521330833435,

```
                                        -0.047524284571409225 ,
                                        0.33808082342147827] ,
                        'link_ang_vel ': 0.88760805 ,
                        'link_vel ': 0.079078786 ,
                        'qpos ': 0.6932341 ,
                        'target_qpos ': 1.5607432007789612}}
reward : 1.132491 { 'object_static ': False , 'open_enough ': False , 'success ': False}
——step #144————
{ 'eval_info ': { 'object_static ': False , 'open_enough ': False , 'success ': False} ,
 'state_info ': { 'gripper_pos ': [ -0.45110204815864563 ,
                                        -0.04662160575389862 ,
                                        0.34376877546310425] ,
                        'link_ang_vel ': 0.7597413 ,
                        'link_vel ': 0.06768684 ,
                        'qpos ': 0.74134094 ,
                        'target_qpos ': 1.5607432007789612}}
reward : 1.180598 { 'object_static ': False , 'open_enough ': False , 'success ': False}
——step #145————
{ 'eval_info ': { 'object_static ': False , 'open_enough ': False , 'success ': False} ,
 'state_info ': { 'gripper_pos ': [ -0.44524049758911133 ,
                                        -0.05206315964460373 ,
                                        0.3193126618862152] ,
                        'link_ang_vel ': 6.543907e-11 ,
                        'link_vel ': 5.830078e-12 ,
                        'qpos ': 0.7430267 ,
                        'target_qpos ': 1.5607432007789612}}
reward : 1.182283 { 'object_static ': False , 'open_enough ': False , 'success ': False}
——step #146————
{ 'eval_info ': { 'object_static ': False , 'open_enough ': False , 'success ': False} ,
 'state_info ': { 'gripper_pos ': [ -0.4403909146785736 ,
                                        -0.05844297260046005 ,
                                        0.3286547064781189] ,
                        'link_ang_vel ': 0.10299364 ,
                        'link_vel ': 0.009175893 ,
                        'qpos ': 0.7472786 ,
                        'target_qpos ': 1.5607432007789612}}
reward : 1.186535 { 'object_static ': False , 'open_enough ': False , 'success ': False}
——step #147————
{ 'eval_info ': { 'object_static ': False , 'open_enough ': False , 'success ': False} ,
 'state_info ': { 'gripper_pos ': [ -0.44815579056739807 ,
                                        -0.0557868666946888 ,
                                        0.3390086591243744] ,
                        'link_ang_vel ': 1.17256285e-08 ,
                        'link_vel ': 1.0446578e-09 ,
                        'qpos ': 0.74716485 ,
                        'target_qpos ': 1.5607432007789612}}
reward : 1.186422 { 'object_static ': False , 'open_enough ': False , 'success ': False}
——step #148————
```

{'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.44925457239151,
                                −0.06130119040608406,
                                0.3416590392589569],
                'link_ang_vel': 1.2060614,
                'link_vel': 0.10745053,
                'qpos': 0.80161244,
                'target_qpos': 1.5607432007789612}}
reward: 1.240869 {'object_static': False, 'open_enough': False, 'success': False}
——step #149————
{'TimeLimit.truncated': True,
 'eval_info': {'object_static': False, 'open_enough': False, 'success': False},
 'state_info': {'gripper_pos': [−0.4398714005947113,
                                −0.054651543498039246,
                                0.3401172459125519],
                'link_ang_vel': 0.45961165,
                'link_vel': 0.040947657,
                'qpos': 0.8030219,
                'target_qpos': 1.5607432007789612}}
reward: 1.242279 {'object_static': False, 'open_enough': False, 'success': False}