

## Appendix

### A. Teacher Reward Formulations

We decompose the door-opening task into six stages: (0) Walk to door, (1) Pre-grasp, (2) Grasp, (3) Open, (4) Swing, and (5) Pass through door. Table 2 summarizes the stage-dependent shaping terms used for the teacher policy.

### B. Synthetic Generation Pipeline of Doors

The procedural generation of doors can be divided into two phases. In phase 1, we generate the physical properties of the doors. In phase 2, we apply randomized texture and lighting.

Table 3 summarizes the physical property randomization ranges. We first spawn the geometries for the wall, door panel, push-bar / handle, floor, and latch. Then we add physical joints for the door hinge and handle. The latch is modeled as a mimic joint attached to the joint angle of the handle. Damping, stiffness, and max force are added to the actuators. The door handle actuator is set to have a -5 degrees (upwards) target joint position to simulate the tension of the spring-loaded handle joint even at level position. Additional random features such as key hole, door frame, and other decorations are spawned each at 50% chance.

We make use of OmniPBR materials and create multiple variants for each by randomizing sub-identifier, texture transform, albedo color, tint color, etc. Every [0.9, 1.1] seconds, a geometry in the scene will have its material randomly drawn. For background dome light texture, we use all publicly available ones in Omniverse, plus an additional 5233 ones from Poly Haven, covering diverse indoor, outdoor, and various times-of-day scenes.

Property	Range	Unit
Panel Width	0.8-1.1	m
Panel Height	1.9-2.2	m
Handle Height	0.85-0.95	m
Handle to Edge Distance	0.04-0.1	m
Handle Type	{knob, lever, pushbar, handle, flat}	
Open Handedness	{left, right}	
Open Direction	{in, out}	
Weight	80-120	kg
Hinge Max Force	20-30	Nm
Hinge Damping	5-10	$(\text{kg m}^2) / (\text{s}^2 \text{ }^\circ)$
Hinge Stiffness	10-20	$(\text{kg m}^2) / (\text{s}^2 \text{ }^\circ)$
Handle Max Force	1-3	Nm
Handle Damping	0.1-0.6	$(\text{kg m}^2) / (\text{s}^2 \text{ }^\circ)$
Handle Stiffness	30-50	$(\text{kg m}^2) / (\text{s}^2 \text{ }^\circ)$

Table 3. Physical property randomization range of doors in IsaacLab.

### C. Teleoperation Baseline Setup

For the experiments in Section 3, we use a PICO 4 Ultra headset with two handheld controllers for both expert and

non-expert teleoperators. The teleoperation interface outputs a command consisting of three upper-body SE(3) poses (head and both wrists), finger joint angles, waist height, and a planar navigation command specifying desired root linear velocity  $\mathbf{v} \in \mathbb{R}^2$  and angular velocity  $\omega \in \mathbb{R}$  for heading control. We employ the Pinocchio library [7] to solve inverse kinematics and map wrist poses to joint-space configurations.

### D. Real World Deployment Setup

We conduct our experiments on a 29-DoF Unitree G1 humanoid robot, equipped with two 7-DoF 3-finger dexterous hands. Perception is provided by an Intel RealSense D435i camera, without the depth output. Policy inference runs on a desktop workstation with an Intel i9-14900K CPU and an NVIDIA RTX 4090 GPU.

Term	Expression	Weight	Stage(s)
<b>Termination / Generic penalties</b>			
Termination	$\mathbb{1}_{\{\text{termination}\}}$	-1000.0	0-5
Delta action rate	$\ \Delta a_t\ _2^2$	-0.01	0-5
DoF velocity	$\ \dot{\mathbf{q}}_{\text{upper, non-finger}}\ _2^2$	$-1.0 \times 10^{-3}$	0-5
DoF acceleration	$\ \ddot{\mathbf{q}}_{\text{upper, non-finger}}\ _2^2$	$-1.0 \times 10^{-5}$	0-5
DoF position limits	$\sum \max(0,  \mathbf{q}_i - \mathbf{q}_{\text{limit}, i} )$	-5.0	0-5
Finger primitive limits	$ \text{clip}(u_{\text{finger}}, [l, u]) - u_{\text{finger}} $	-1.0	0-5
Humanly DoF limit	$\sum (\text{clip}(\mathbf{q} - \mathbf{q}_{\text{lower}}, \max = 0) + \text{clip}(\mathbf{q} - \mathbf{q}_{\text{upper}}, \min = 0))$	-1.0	0-5
DoF overspeed	$\sum \max(0,  \dot{\mathbf{q}}_i  - 2.0)^2$	-0.1	0-5
Undesired contact	$\sum \mathbb{1}_{\{\ \mathbf{f}_{\text{contact}, i}\  > 1\}}$	-0.2	0-5
Door frame contact	$\sum \ \mathbf{f}_{\text{door frame}}\ _2$	-0.1	0-5
Door panel contact	$\sum \ \mathbf{f}_{\text{door panel}}\ _2$	-0.1	0-5
Upright penalty	$\ R_{\text{torso}}[0, 0, 1]^\top - [0, 0, 1]^\top\ _2^2$	-1.0	0-5
HOMIE action limit	$\sum \max(0,  u_{\text{homie}, i}  - 1.0)$	-1.0	0-5
<b>Stage 0: Walk to door</b>			
Walk to door	$\exp(-\ \mathbf{v}_{\text{robot}} - v_{\text{target}} \hat{\mathbf{d}}_{\text{door}}\ _2^2 / (2 \cdot 0.15^2)), \sigma = 0.15$	5.0	0
Upper body deviation	$\ \mathbf{q}_{\text{upper, non-finger}} - \mathbf{q}_{\text{resting}}\ _1$	-1.0	0, 5
Face door	$ \text{wrap}_\pi(\ \text{axis-angle}(R_{\text{door}})\ _2) $	-1.0	0-2, 5
<b>Stage 1: Pre-grasp</b>			
Hand-handle orientation	$\exp(-\ \text{wrap}_\pi(\ \text{axis-angle}(R_{\text{hand}} R_{\pm 90})\ _2)\ ^2 / (2 \cdot 0.6^2))$	3.0	1-4
Pregrasp finger pose	$\text{track}(\mathbf{q}_{\text{finger}}, \mathbf{q}_{p0}, \sigma_{\text{pos}} = 0.3) + \text{track}(\dot{\mathbf{q}}_{\text{finger}}, 0.6, \sigma_{\text{vel}} = 0.2)$	1.5	0-1, 5
Unused arm deviation	$\ \mathbf{q}_{\text{unused arm}} - \mathbf{q}_{\text{rest}}\ _1$	-1.0	1-4
Pre-grasp target distance	$\text{track}(\ \mathbf{p}_{\text{hand}} - \mathbf{p}_{\text{pre-grasp}}\ , 0, \sigma = 0.2) + \text{track}(\ \mathbf{v}_{\text{hand}} - v_{\text{target}} \hat{\mathbf{d}}\ , 0, \sigma = 0.15)$	6.0	1
Penalty not standing still	$\ \mathbf{u}_{\text{HOMIE}, [0:3]}\ _2$	-15.0	1-3
<b>Stage 2: Grasp</b>			
Grasp finger DoF pose	$\text{track}(\mathbf{q}_{\text{finger}}, \mathbf{q}_{p1}, \sigma_{\text{pos}} = 0.3) + \text{track}(\dot{\mathbf{q}}_{\text{finger}}, 0.6, \sigma_{\text{vel}} = 0.2)$	3.0	2-4
Grasp target distance	$\exp(-\ \mathbf{p}_{\text{hand}} - \mathbf{p}_{\text{grasp}}\ _2^2 / (2 \cdot 0.1^2))$	3.0	2-4
Grasp force	$\sum (- f_{\text{palm}, y, z}  + f_{\text{palm}, x})$	0.2	1-4
<b>Stage 3: Open door</b>			
Push door handle	$\dot{\theta}_{\text{handle}} + \text{clip}(\theta_{\text{handle}}, 0, 45^\circ) / 45^\circ$	6.0	3
Push door hinge	$10\dot{\theta}_{\text{hinge}} + \text{clip}(\theta_{\text{hinge}}, 0, 90^\circ) / 90^\circ$	6.0	3-4
Push door force	$\text{clip}(f_{\text{hand}, x}, 0, 20)$	0.3	3
<b>Stage 4: Swing door &amp; Stage 5: Pass through Door</b>			
Don't push door handle	$-\dot{\theta}_{\text{handle}} + (45^\circ - \theta_{\text{handle}}) / 45^\circ$	3.0	4-5
Target root distance	$\text{track}(\mathbf{v}_{\text{root}} \cdot \hat{\mathbf{d}}_{\text{target}}, v_{\text{target}}, \sigma = 0.2) + \text{track}(\ \mathbf{p}_{\text{root}} - \mathbf{p}_{\text{target}}\ , 0, \sigma = 0.2)$	12.0	4-5
Penalty standing still	$\exp(-\ \mathbf{u}_{\text{HOMIE}, [0:3]}\ _2^2 / (2 \cdot 0.05^2))$	-1.0	4
<b>Always-on rewards</b>			
Stage progress	$\text{stage}_{\text{current}}$	1.0	0-5
Task completion	$\mathbb{1}_{\{\text{complete}\}}$	4.0	0-5
Success save time	$\mathbb{1}_{\{\text{success}\}} \cdot \text{remaining time ratio}$	0.5	0-5

Table 2. Reward components for door opening task.  $\text{Track}(x, \mu, \sigma)$  denotes Gaussian tracking reward  $\exp(-(x - \mu)^2 / (2\sigma^2))$ .