A Implementation Details

Our implementation uses PyTorch for both training and inference. For point cloud processing, we employ different sampling strategies across model components. The DDPM input processing randomly samples 2048 points from the original 65536-point object cloud. For the Embodiment-Agnostic Grasp Refinement stage, we utilize either importance sampling based on contact map values to select 256 points followed by FPS sampling for another 256 points when using anchor points, or direct random sampling of 512 points otherwise.

B Human-to-Robot Retargeting

We implement the retargeting process using the AnyTeleop [28] framework with hand-specific scaling factors. For the Shadowhand, we maintain the original scale, while the Allegro and LeapHand configurations use scaling factors of 1.9 and 1.8 respectively. To accommodate AnyTeleop's sequence-based design, we replicate each static grasp frame 20 times.

The grasp optimization process employs modified BoDex parameters to preserve the initial retargeted configurations while ensuring stability. We utilize reduced joint angle search amplitudes of 0.01 for Shadowhand and 0.05 for LeapHand and Allegro, balancing optimization efficacy with position maintenance.

C Additional Results

C.1 Cross-Embodiment Performance

Table 2 presents the detailed performance metrics across different robotic hands. The performance variation between different hand types can be attributed to the scaling effects in the retargeting process. While necessary for adaptation, the scaling of objects for Allegro and Leaphand (1.9 and 1.8 respectively) introduces additional challenges in maintaining grasp stability, resulting in lower success rates compared to the unscaled ShadowHand configuration.

Table 2: Cross-embodiment performance comparison

Model Configuration	Success Rate	Chamfer Distance
Ours Allegro (3 hand version)	65%	5.8
Ours Leaphand (3 hand version)	40%	6.5