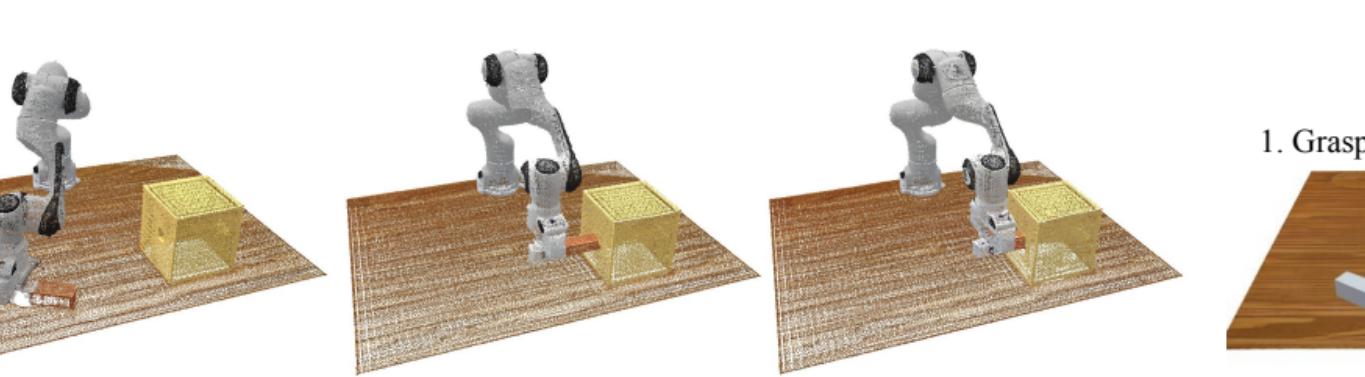
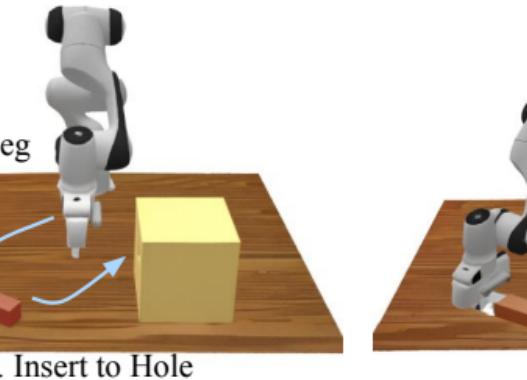


3D latents from
multi-view observations



Predictive rollout of future 3D latent states



1. Grasp Peg
2. Insert to Hole

Robotic action execution

