x_1

x_2

Encoder z_1 Decoder x_2

1. Latent Action Discretization

Reach to
the cup

LatentVLA z_1

2. Latent Pretraining

Pick up the milk and
put it in the sink

LatentVLA a_1

Latent Action Pretraining ⟶ **Action Finetuning**