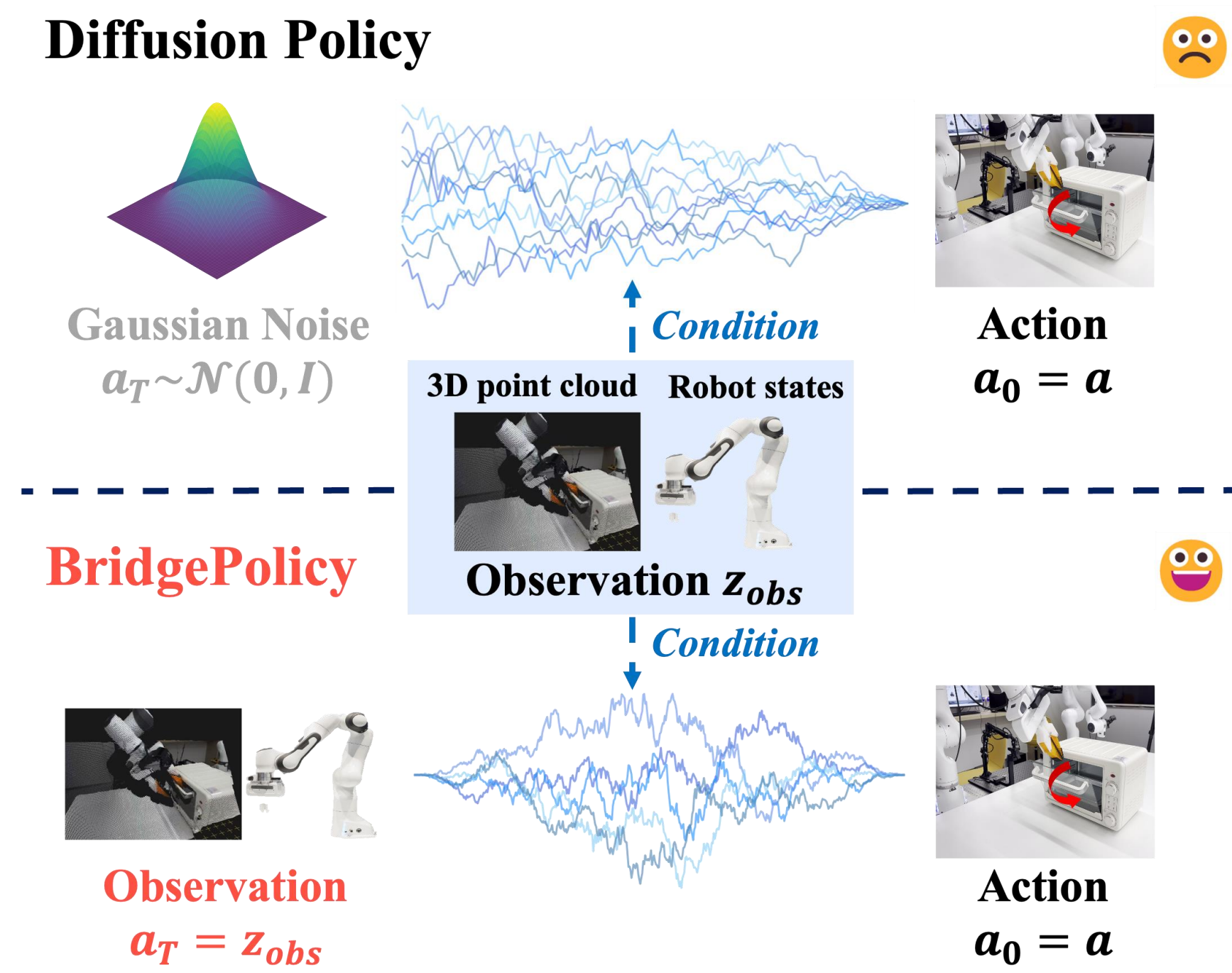




# Sample from What You See: Visuomotor Policy Learning via Diffusion Bridge with Observation-Embedded Stochastic Differential Equation

## Motivation:

Traditional diffusion-based policy underutilized the observation information by only treating them as high-level condition during the sampling 🙄 process, leading to suboptimal performance.



## Our Solution:

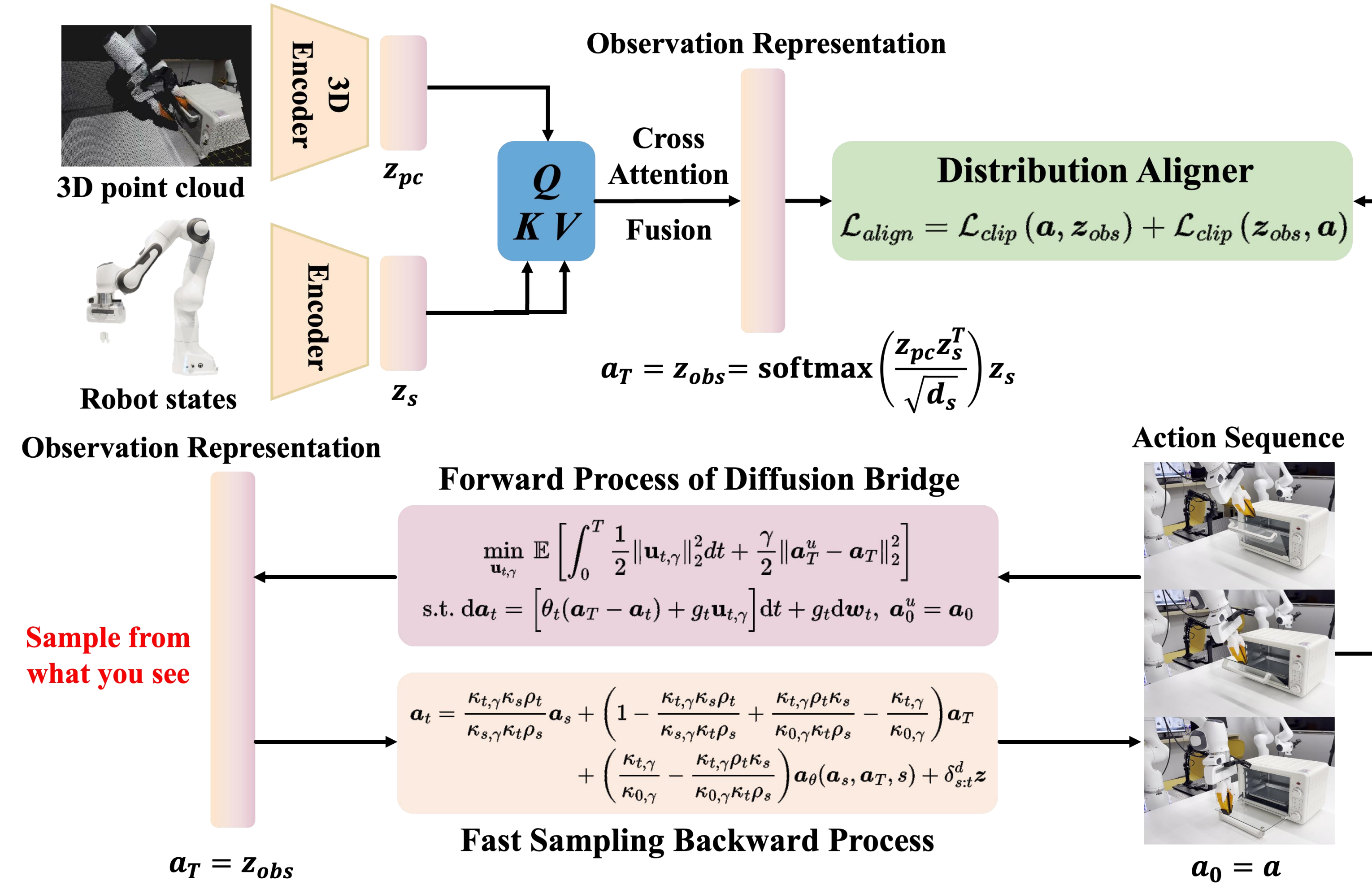
By constructing an observation-informed diffusion SDE trajectory via **Diffusion Bridge**, **BridgePolicy** enables sampling to start from a rich, informative prior rather the uninformative random noise.

## Challenges:

- Diffusion Bridge can only connect one-to-one distributions while the observations are inherently multi-modal. 🤔
- Standard Diffusion Bridge require the connected distributions share the same data shape, which is barely satisfied in robotic observations.

## Method:

To this end, we propose **BridgePolicy**, a novel generative policy that embed the observation into the diffusion SDE trajectory via the formulation of diffusion bridge, therefore directly samples actions from observation-informed prior rather than random noise.



## Simulation tasks

Methods\Task	MetaWorld Easy	MetaWorld Medium	MetaWorld Hard	MetaWorld Very Hard	DexArt	Adroit	Average
DP [5]	0.79±0.27	0.31±0.24	0.10±0.12	0.26±0.25	0.45±0.08	0.31±0.17	0.37
DP3 [55]	0.87±0.20	0.61±0.29	0.40±0.37	0.51±0.33	0.57±0.08	0.68±0.06	0.60
Simple DP3 [55]	0.86±0.12	0.59±0.22	0.38±0.26	0.47±0.28	0.48±0.07	0.68±0.05	0.58
FlowPolicy [56]	0.86±0.16	0.67±0.21	<b>0.59±0.30</b>	0.76±0.26	0.54±0.09	0.70±0.12	0.68
<b>BridgePolicy (Ours)</b>	<b>0.91±0.24</b>	<b>0.75±0.28</b>	0.58±0.31	<b>0.79±0.32</b>	<b>0.60±0.07</b>	<b>0.81±0.12</b>	<b>0.74</b>

## Quantitative Results:

Achieve state-of-the-art performance on 52 simulation task across 3 benchmarks and 5 real-world tasks.

## Real-world tasks

	Oven-Closing	Oven-Opening	Pick Place	Pour	Unplug	Average
Simple DP3	0.8	0.6	0.6	0.6	0.7	0.66
DP3	0.9	0.9	0.7	0.6	0.7	0.76
FlowPolicy	<b>1.0</b>	0.7	0.5	0.1	0.5	0.56
<b>BridgePolicy</b>	<b>1.0</b>	<b>1.0</b>	<b>0.8</b>	<b>0.8</b>	<b>0.9</b>	<b>0.90</b>

## Qualitative Results:

Demonstrate more stable and controllable behaviors than the Diffusion-based and Flow-based policies when performing the real-world tasks.

