From Diffusion to Flow: Fast and Coherent Policy Distillation via Probability Flow ODEs

Jing Wang[1], Xihua Wang[1], and Li Cheng[1]
[1] Electrical and Computer Engineering, University of Alberta, Edmonton, Canada.
Email: mailto:jing39@ualberta.ca

INTRODUCTION

Diffusion Policies (DPs) [1] have achieved state-of-the-art performance in visuomotor control by modeling multimodal action distributions through iterative denoising. However, inference requires hundreds of sampling steps, making real-time deployment impractical. The existing distillation methods compress DPs into single-step students but typically lose trajectory-level stochasticity and temporal coherence, which is key to DPs' generative power.

MATERIALS AND METHODS

This paper introduces One-Step Flow DP (OneFlow-DP), a novel distillation framework that reformulates the stochastic diffusion process into a deterministic probability flow ODE (PF-ODE) [2]. Instead of endpoint matching, DGF aligns the student's velocity field with the teacher's score-driven dynamics across intermediate trajectories. This score-alignment ensures that different initial noises are transported along distinct, coherent paths, preserving multimodality and uncertainty structure.

IMPLEMENTATION

The student is trained via a flow-matching loss (global score alignment) plus an auxiliary action prediction loss (local fidelity at terminal step). This enables one-step inference while retaining diversity and coherence. Training converges within around 20 epochs.

RESULTS AND DISCUSSION

We evaluate DGF on Push-T benchmark, comparing with DP [1], Consistency Policy (CP) [3], OneDP [4], and Variational Diffusion Distillation (VDD) [5].

Efficiency Gains: DGF reduces inference time from around 160 ms (DP) to 2 ms (~80× speedup) while also lowering FLOPs by over 100×.

Performance Gains: On the challenging Push-T task, DGF is the only distillation method that enhances the DP teacher's performance, reaching a coverage score of 0.92 versus 0.86 for DP, whereas all other distillation approaches reduce performance.

Discussion: Unlike endpoint-matching methods, DGF distills not just outputs but the generative dynamics of diffusion teachers. Different noise seeds lead to coherent but distinct trajectories to preserve the stochasticity despite deterministic inference.

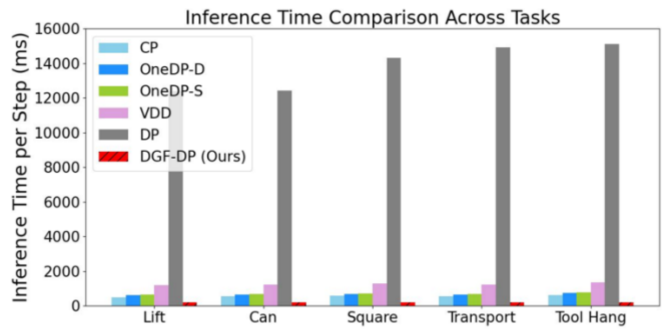

Fig 1 Inference time comparison across *Robomimic* benchmark, demonstrating that our DGF-DP achieves the fastest inference.

CONCLUSIONS

OneFlow-DP bridges the gap between expressive but slow diffusion policies and fast real-time control. It offers practical deployment potential for robotics in manufacturing, logistics, and assistive domains. Thus, our work provides a principled and theoretically grounded approach to distilling diffusion policies. By casting distillation as score-aligned flow matching, it preserves generative dynamics, stochasticity in a single-step model, advancing real-time robot policy learning.

REFERENCES

[1] Chi C et al. Int J Robot Res 42: 1150-1165, 2023.
[2] Song Y et al. Int Conf Learn Represent: 1-12, 2021.
[3] Prasad A et al. Robot Sci Syst: 1-8, 2024.
[4] Wang Z et al. Arxiv, 2024
[5] Zhou H et al. Adv Neural Inf Process Syst 36: 1-12, 2024.

Table 1: Performance on Push-T. Inference cost is measured in per evaluation step latency (ms) and estimated FLOPs per trajectory.

Method	Push-T Coverage Score (↑)	Inference Time (ms) (↓)	FLOPs ($\times 10^9$) (↓)
CP [3]	0.84 ± 0.04	4.59 ± 0.9	6.7 ± 0.2
OneDP [4]	0.80 ± 0.06	6.06 ± 0.9	9.1 ± 0.3
VDD [5]	0.85 ± 1.05	12.53 ± 1.2	18.5 ± 0.6
DP [1]	0.86 ± 0.04	159.6 ± 5.5	254.9 ± 4.1
OneFlow-DP (Ours)	**0.92 ± 0.03**	**2.0 ± 0.8**	**2.5 ± 0.1**