2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

SeFA-Policy: Fast and Accurate Visuomotor Policy Learning with Selective Flow Alignment

Rong Xue*, Jiageng Mao*, Mingtong Zhang, Yue Wang
University of Southern California *Equal Contribution
SeFA-Policy

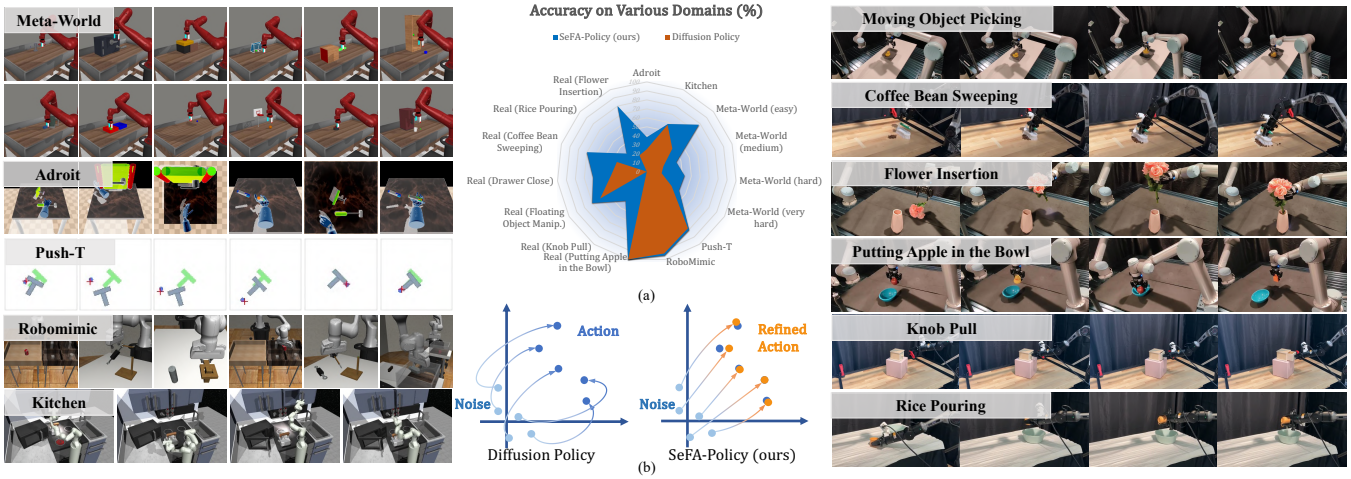


Fig. 1: **Selective Flow Alignment** (**SeFA**) is a visual imitation learning algorithm that utilizes rectified flow with selective alignment, achieving superior effectiveness in diverse simulation and real-world tasks, with a significant inference acceleration. (a) Accuracy on various domains. (b) Sampling flow from noise to action of SeFA and Diffusion Policy.

***Abstract*— Developing efficient and accurate visuomotor policies poses a central challenge in robotic imitation learning. While recent rectified flow approaches have advanced visuomotor policy learning, they suffer from a key limitation: After iterative distillation, generated actions may deviate from the ground-truth actions corresponding to the current visual observation, leading to accumulated error as the reflow process repeats and unstable task execution. We present Selective Flow Alignment (SeFA), an efficient and accurate visuomotor policy learning framework. SeFA resolves this challenge by a selective flow alignment strategy, which leverages expert demonstrations to selectively correct generated actions and restore consistency with observations, while preserving multimodality. This design introduces a consistency correction mechanism that ensures generated actions remain observation-aligned without sacrificing the efficiency of one-step flow inference. Extensive experiments across both simulated and real-world manipulation tasks show that SeFA surpasses state-of-the-art diffusion-based and flow-based policies, achieving superior accuracy and robustness while reducing inference latency by over 98%. By unifying rectified flow efficiency with observation-consistent action generation, SeFA provides a scalable and dependable solution for real-time visuomotor policy learning. Code is available on SeFA code.**

I. Introduction

Imitation learning relies on accurate action predictions and fast inference to successfully perform complicated real-world tasks. Generative modeling techniques such as Diffusion Policy [1] have recently achieved strong performance in complex manipulation tasks, but their reliance on multi-step iterative denoising makes them computationally expensive and unsuitable for real-time control. Flow-based models have emerged as a promising alternative, enabling fewer-step action generation [2] and rectification [3] by transporting nearly straight from noise to action space, thus significantly reducing inference latency.

Despite the efficiency, few-step sampling introduces discretization error, and rectification introduces inconsistency between observation and action during distillation. When applying rectification, the reflow policy [3] is trained upon the noise-action pairs generated by a well-trained policy. However, the generated actions are not the same as ground-truth actions, i.e., the generated action and visual observation pair could not exactly match. In terms of diffusion-based models, the predicted action is distinct from the actions reflected in the visual condition. Such inconsistency might be tolerable in image generation where perceptual similarity suffices, but in robotic control, even minor inconsistencies between observations and actions can accumulate and lead to task failures. This distillation-induced inconsistency therefore represents a fundamental barrier to deploying flow-based policies in effective real-time visuomotor control.

To overcome this limitation, we propose **Selective Flow Alignment** (**SeFA**), a flow-based visuomotor policy with a selective alignment strategy. The straight paths in flow-based

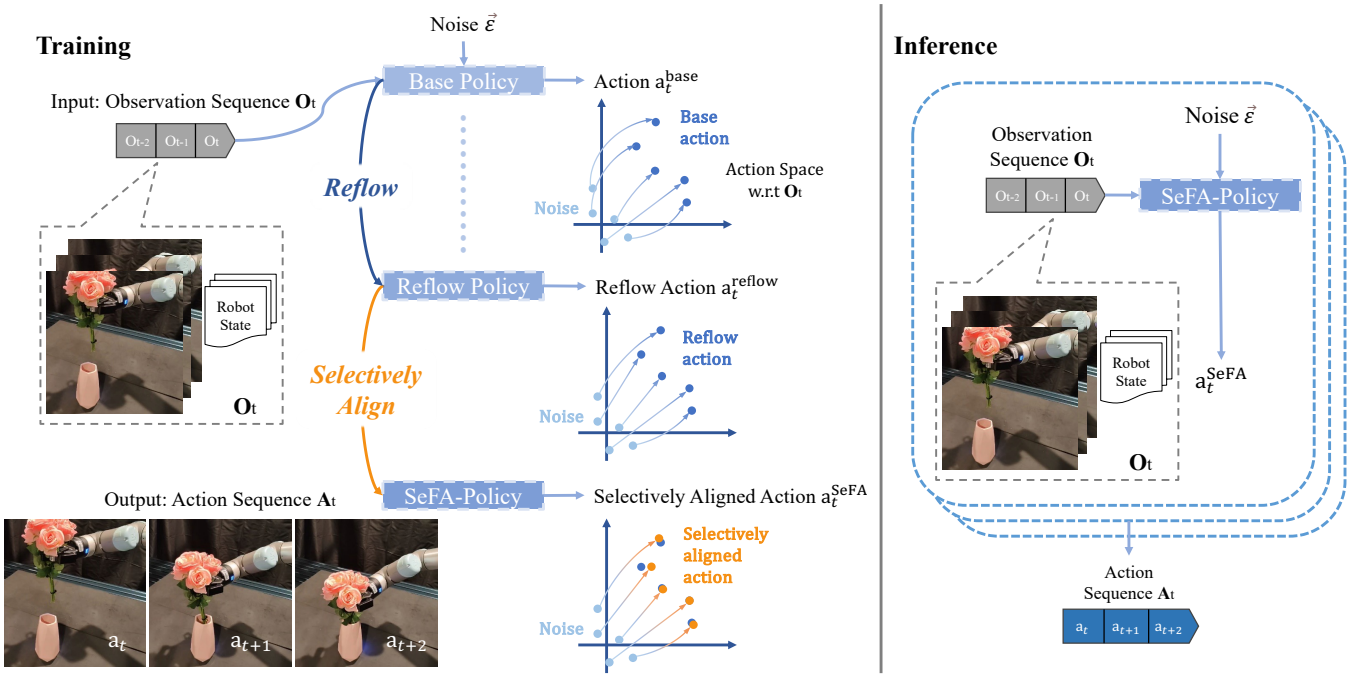


Fig. 2: **Overview of SeFA.** We train a visuomotor policy in an iterative manner to transport straight between noise distribution and target action space, hence enabling lightning one-step sampling during inference. The action flow is selectively *aligned* with observations, lowering the potential accumulated error brought by multiple reflows.

models are computationally efficient because they can be sampled in a few or even one step. However, the straightening process [3] often accumulates errors from the base model, which leads to the observation–action inconsistency. SeFA leverages expert demonstrations to align the sampling paths with observations while maintaining the straightness of the paths. Crucially, this alignment is applied in a selective manner, preserving action diversity and multimodality while eliminating harmful mismatches. By combining efficiency in straight sampling paths with observation-consistent alignment, SeFA enables one-step action synthesis that is both fast and reliable for real-time visuomotor control.

We summarize our main contributions as follows:

- Selective Flow Alignment (SeFA): We introduce a novel visuomotor policy learning framework that improves flow-based models with a selective alignment strategy to address the observation–action inconsistency problem.
- Consistency Correction with Efficiency: Our method achieves one-step action generation that is both observation-aligned and computationally efficient, preserving efficiency while ensuring robustness in control.
- Extensive Evaluation: We demonstrate SeFA 's effectiveness across diverse simulated and real-world robotic tasks, where it consistently surpasses diffusion-based and flow-based baselines in accuracy, robustness, and inference speed.

II. Related Work

Diffusion Models in Robotics. Diffusion models excel at representing complex multimodal distributions with stable training dynamics and hyperparameter robustness, gaining traction across robotics applications including motion planning [4], [5], [6], [7], [8], imitation learning [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], goal-conditioned imitation learning [20], [21], [22], [23], and grasp prediction [24]. However, their iterative denoising process renders them impractical for real-time robotics applications. Our proposed SeFA overcomes this limitation through rectified action flow, enabling deterministic, fast action synthesis without compromising accuracy.

Recent work has explored flow matching [25], a diffusion variant, for representing complex continuous action distributions. AdaFlow [2] introduces a variance-adaptive ODE solver with adjustable step sizes, but only achieves one-step generation for uni-modal distributions due to lacking reflow, and hasn't been validated on real robots or comprehensive simulation domains. π_0 [26] leverages pre-trained Vision-Language Models with flow matching for high-frequency action generation, but still requires 10 integration steps. In contrast, our method maintains high precision even with one-step prediction.

Accelerating Diffusion Models for Robotics. Efforts to speed up diffusion models have been explored extensively in both image generation [27], [28], [29], [30] and robotics. However, these methods often require complex distillation processes or introduce constraints, such as overly smooth trajectories in Dynamical Motion Primitives (DMPs) [31]. Streaming Diffusion Policy (SDP) [32] and related approaches like Rolling Diffusion [33] and Temporally Entangled Diffusion [34] improve speed through parallelization or buffering, but they often incur significant memory overhead or require intricate implementation.

Overall, while significant progress has been made in

accelerating diffusion models for robotics, achieving real-time performance without sacrificing accuracy remains a challenge. Our work builds on these advancements by introducing SeFA, which leverages rectified flow to achieve fast and accurate visuomotor control and introduces a selective *alignment* strategy to mitigate potential accumulated errors from multiple reflows. By replacing the iterative denoising process with a deterministic coupling, SeFA offers a streamlined and efficient solution for real-time robotic applications.

III. Method

A visuomotor policy solves the task of observing a sequence of visual observations and predicts the next action to execute in the environment. We formulate SeFA-Policy as a flow-based model in §III-A. Then, we introduce our proposed *Selective Flow Alignment* strategy to alleviate performance degradation and improve the performance in §III-B.

A. Flow-based Model

We begin by formalizing our approach to action space modeling. The fundamental objective is to establish a bijective mapping between standard Gaussian noise $\mathbf{a}_T \in \mathbb{R}^d$ and the target action distribution $\mathbf{a}_0 \in \mathbb{R}^d$ given the visual observations $\mathbf{O}$. During the initial training phase, we define a drift vector field $v : \mathbb{R}^d \to \mathbb{R}^d$ that guides the flow along trajectories approximating the direct linear path from $\mathbf{a}_T$ to $\mathbf{a}_0$. This is achieved by minimizing the expected squared deviation between the drift and the ideal direction $(\mathbf{a}_0 - \mathbf{a}_T)$ through the following least squares optimization problem:

$$\min_v \int_0^1 \mathbb{E}\left[||(\mathbf{a}_0 - \mathbf{a}_T) - v\big(\mathbf{a}_t, t, \mathbf{O}\big)||^2\right] \mathrm{d}t, \qquad (1)$$

where $\mathbf{a}_t$ is the linear interpolation of $\mathbf{a}_T$ and $\mathbf{a}_0$, i.e., $\mathbf{a}_t = \frac{t}{T}\mathbf{a}_T + \frac{T-t}{T}\mathbf{a}_0$, where $t \in [0, T]$. In the following, we will omit the conditioning on visual observations $\mathbf{O}$ for simplicity. Naturally, $\mathbf{a}_t$ follows the ODE of $\mathrm{d}\mathbf{a}_t = (\mathbf{a}_0 - \mathbf{a}_T)\mathrm{d}t$, where any update of $\mathbf{a}_t$ requires the information of the target clean action $\mathbf{a}_0$. By fitting the drift v with $\mathbf{a}_0 - \mathbf{a}_T$, the action flow causalizes the paths of linear interpolation $\mathbf{a}_t$, relieving the burden of involving the target action (which is unknown during inference) when simulating the ODE flow.

The solution to Equation (1) constitutes our base policy network, denoted as v^{base}. When provided with a noise sample $\mathbf{a}_T$, the base policy generates the corresponding action $\mathbf{a}_0$.

To make a distinction, v^{base} is trained with the random noise and the groundtruth action pairs $(\mathbf{a}_T, \mathbf{a}_0)$, and the derived coupling is denoted as $(\mathbf{a}_T^{\text{base}}, \mathbf{a}_0^{\text{base}})$. The action flow induced between $\mathcal{N}(0, \mathbf{I})$ and $\mathcal{A}$ is:

$$\mathrm{d}\mathbf{a}_t^{\text{base}} = v^{\text{base}}(\mathbf{a}_t^{\text{base}}, t)\mathrm{d}t, \quad t \in [0, T], \qquad (2)$$

which converts the noise $\mathbf{a}_T^{\text{base}} \in \mathcal{N}(0, \mathbf{I})$ in the coupling $(\mathbf{a}_T^{\text{base}}, \mathbf{a}_0^{\text{base}})$ to the action $\mathbf{a}_0^{\text{base}}$ which follows the conditioned expert action distribution.

Reflow. After the drift v^{base} is estimated, we can accelerate inference by training a reflow policy [3] that straightens the sampling paths. In our settings, we first randomly sample noise $\mathbf{a}_T^{\text{base}} \sim \mathcal{N}(0, \mathbf{I})$ and then forward generate $\mathbf{a}_0^{\text{base}}$ following Equation (2). This coupling $(\mathbf{a}_T^{\text{base}}, \mathbf{a}_0^{\text{base}})$ ensures optimal transport efficiency—specifically, it guarantees that the transport cost remains lower than that of any arbitrary (action×noise) pairing across all convex cost functions, a property that follows directly from Jensen's inequality. Through this formulation, we establish a principled deterministic mapping between the action space $\mathcal{A}$ and the standard Gaussian distribution $\mathcal{N}(0, \mathbf{I})$.

Following [3], we train the reflow policy network using the optimal coupling $(\mathbf{a}_T^{\text{base}}, \mathbf{a}_0^{\text{base}})$ as the substitution of the former pairs. The reflowed action flow v^{reflow} satisfies:

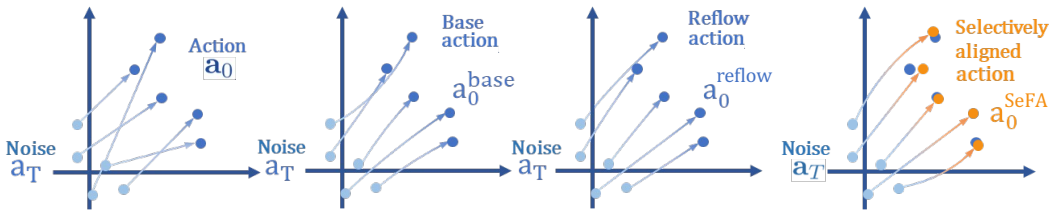
$$\mathrm{d}\mathbf{a}_t^{\text{reflow}} = v^{\text{reflow}}(\mathbf{a}_t^{\text{reflow}}, t)\mathrm{d}t, \quad t \in [0, T], \qquad (3)$$

where $\mathbf{a}_t^{\text{reflow}}$ is the linear interpolation of $\mathbf{a}_T^{\text{base}}$ and $\mathbf{a}_0^{\text{base}}$, i.e., $\mathbf{a}_t^{\text{reflow}} = \frac{t}{T}\mathbf{a}_T^{\text{base}} + \frac{T-t}{T}\mathbf{a}_0^{\text{base}}$, where $t \in [0, T]$. This procedure straightens the paths of the flows. The straighter the paths are, the smaller the time-discretization error in numerical simulation will be. Perfectly straight paths can be exactly simulated with a single Euler step. This addresses the very bottleneck of high inference cost in existing continuous-time ODE-based models, such as the Diffusion Policy [1] built upon Probability Flow ODE [35].

B. Selective Flow Alignment

It is noteworthy that v^{reflow} is trained upon the couplings generated by v^{base}. However, the generated actions $\mathbf{a}_0^{\text{reflow}}$ are not always the same as ground-truth actions, i.e., the generated action and visual observation pair could not exactly match. In terms of diffusion-based models, the generated action could be distinct from the actions reflected in the visual condition. Therefore, although reflow process can accelerate inference, it introduces inconsistency between observation and action. Unlike image generation where minor output variations are acceptable or sometimes even desired, even slight discrepancies in the generated actions can lead to task failures in robot learning, and the error could accumulate as the reflow procedure repeats.

To address this performance degradation introduced by reflow, we introduce a selective flow alignment strategy. For each generated action $\mathbf{a}_0^{\text{reflow}}$, we methodically search the ground-truth action dataset to identify its nearest neighbor. When the Euclidean distance between these actions falls below a predetermined threshold ϵ, we replace the generated action with its ground-truth counterpart. We select ϵ heuristically to balance refinement capability and efficiency: too small values limit available expert actions for refinement, while excessively large values increase computational cost without improving selection quality. Since action spaces are conditioning-dependent, this approach maintains fixed conditioning while identifying the optimal coupling. This strategy effectively addresses two critical scenarios: (1) When a suitable nearest neighbor exists within the threshold, the prediction error attributable to the reflow process is rectified, thereby preserving the least transport costs property; and (2) When the reflow model generates a viable alternative solution that significantly differs from the expert policy, we



(a) Random sample pairs (b) Couplings generated via base policy (c) Couplings generated via reflow policy (d) Selectively aligned couplings

Fig. 3: **Sampling trajectories of SeFA-Policy at different stages.** Randomly sampled pairs in (a) have crossing flows. *Couplings* in (b) have been rewired so they do not intersect with each other at the same denoising timestep. The trajectories in (c) and (d) are nearly straight.

intentionally preserve this inherent multimodality to maintain the richness of the action distribution.

After the selective alignment, we obtain a new coupling $(\mathbf{a}_T^{\text{SeFA}}, \mathbf{a}_0^{\text{SeFA}})$. This coupling should maintain the property of optimal transport efficiency, while the actions $\mathbf{a}_0^{\text{SeFA}}$ are aligned with expert demonstrations. We can train the SeFA-Policy network using $(\mathbf{a}_T^{\text{SeFA}}, \mathbf{a}_0^{\text{SeFA}})$:

$$\mathrm{d}\mathbf{a}_t^{\text{SeFA}} = v^{\text{SeFA}}(\mathbf{a}_t^{\text{SeFA}}, t)\mathrm{d}t, \quad t \in [0, T], \tag{4}$$

where $\mathbf{a}_t^{\text{SeFA}}$ is the linear interpolation of $\mathbf{a}_T^{\text{SeFA}}$ and $\mathbf{a}_0^{\text{SeFA}}$, i.e., $\mathbf{a}_t^{\text{SeFA}} = \frac{t}{T}\mathbf{a}_T^{\text{SeFA}} + \frac{T-t}{T}\mathbf{a}_0^{\text{SeFA}}$, where $t \in [0, T]$. The drift should follow the ODE of $\mathrm{d}\mathbf{a}_t^{\text{SeFA}} = (\mathbf{a}_0^{\text{SeFA}} - \mathbf{a}_T^{\text{SeFA}})\mathrm{d}t$.

IV. Simulation Experiments

We systematically evaluate SeFA-Policy on 66 tasks from 5 benchmarks in simulation, including Adroit [36], RoboMimic [37], Meta-World [38], Franka Kitchen [39], and Push-T [40].

A. Algorithm

We first elaborate the training procedure of reflow in Algorithm 1 and our *Selective Flow Alignment* in Algorithm 2.

Algorithm 1: *Reflow* and *coupling* generation

Input: Base policy with velocity estimation v^{base}.
Number of *couplings* N.
Procedure:
// Coupling generation
for $i = 1$ *to* N **do**
  Sample noise $\mathbf{a}_T^{\text{base}} \sim \mathcal{N}(0, \mathbf{I})$.
  Generate action $\mathbf{a}_0^{\text{base}}$ following $\mathrm{d}\mathbf{a}_t = v^{\text{base}}(\mathbf{a}_t^{\text{base}}, t)\mathrm{d}t$ starting from noise $\mathbf{a}_T^{\text{base}}$.
  Construct *coupling* $(\mathbf{a}_T^{\text{base}}, \mathbf{a}_0^{\text{base}})$.
end
// Training
Initialize parameters of $v^{\text{reflow}} \doteq v^{\text{base}}$.
while *terminal condition* **do**
  Sample timestep $t \sim \text{Uniform}([0, 1])$.
  Compute $\mathbf{a}_t^{\text{base}} = t\mathbf{a}_T^{\text{base}} + (T - t)\mathbf{a}_0^{\text{base}}$.
  Evaluate $\mathbb{E}\big[||\mathbf{a}_0^{\text{base}} - \mathbf{a}_T^{\text{base}} - v^{\text{reflow}}(\mathbf{a}_t^{\text{base}},\ t)||^2\big]$.
  Update parameters of v^{reflow}.
end
Output: Reflow policy. *Coupling* $(\mathbf{a}_T^{\text{reflow}}, \mathbf{a}_0^{\text{reflow}})$.

B. Effectiveness

Comparison with State-of-the-Art Methods. We benchmark SeFA-Policy against two leading baselines: Diffusion Policy [41] and the flow-based AdaFlow Policy [2]. To ensure a rigorous and fair comparison, all experimental settings, including training epochs, random seeds, learning rate schedules, and image resolutions, are held constant across all methods. A summary of the results is presented in Table I. We deliberately avoided tuning hyperparameters for each task to maintain a fair experimental protocol. We observe that SeFA-Policy achieves a success rate exceeding 80% on 31 tasks, whereas Diffusion Policy surpasses this threshold on only 16 tasks. The average success rate of SeFA-Policy attains **62.3**%, substantially outperforming Diffusion Policy, which achieves only 38.9%. The policies trained on Franka Kitchen are only given low-dimensional conditioning to test our method without visual inputs. Even under these conditions, SeFA-Policy attains 100% accuracy across all tasks, outperforming the state-of-the-art baseline and demonstrating strong robustness to the modality of conditioning.

The superior performance of our method compared to Diffusion Policy can be attributed primarily to the nature of the sampling trajectories induced by the underlying integration schedule. Diffusion Policy often generated unstable, oscillatory motions and unnecessary deviations from the optimal path. This observation is consistent with prior findings [41], which highlight that the integration schedule must be well-matched to the underlying data distribution. In robotic control tasks, where action distributions are typically concentrated, flow matching schedules facilitate straighter sampling trajectories and reduce integration error. By capitalizing on this property, our approach achieves greater accuracy and stability in robotic control compared to variance-preserving diffusion schedules. Compared to AdaFlow, the state-of-the-art flow-based policy, our method reduces the discretization error by introducing the *SeFA* strategy. Besides, AdaFlow utilizes adjustable integration steps that are typically greater than 1, resulting in lower efficiency compared to our one-step prediction approach.

Comparison with more baselines. We also include Consistency Policy [42] as our baseline. Due to the time-

Algorithm 2: *Selective Flow Alignment* (*SeFA*)

Input: Target action $\mathbf{a}_0^{\text{reflow}}$ and corresponding visual observation sequence $\mathbf{O}$. Action distance threshold δ.

Procedure:

Find the nearest ground-truth condition to $\mathbf{O}$, with its corresponding action $\mathbf{a}_0^{\text{reflow},*}$.

if $\|\mathbf{a}_0^{reflow}, \mathbf{a}_0^{reflow,*}\| < \delta$ **then**

  $\mathbf{a}_0^{\text{SeFA}} \doteq \mathbf{a}_0^{\text{reflow},*}$

else

  $\mathbf{a}_0^{\text{SeFA}} \doteq \mathbf{a}_0^{\text{reflow}}$

end

Output: Selectively Aligned target action $\mathbf{a}_0^{\text{SeFA}}$.

consuming procedure of training a teacher model and then distilling it into a student model in Consistency Policy pipeline, we only evaluate its performance on a few randomly selected tasks across various domains. In RoboMimic tasks, we follow the original settings in its paper. For Push-T and Meta-World tasks, we run its pipeline using our settings. The results are reported in Table II. SeFA-Policy shows consistent improvement on all benchmarks.

Ablation study on reflow and *SeFA*. Some works argue that although the distillation of the diffusion process can be used to accelerate policy synthesis, it is computationally expensive and can hurt both the accuracy and diversity of synthesized actions. In this paper, we show that although this statement holds true in some circumstances, we can still neutralize the drawbacks of reflow by our *SeFA* method. Here we test the performance of base policy, reflow policy, and SeFA-Policy. We choose 3 tasks with significantly different base policy accuracy on Meta-World and Adroit to prove our point. As presented in Table IV, on tasks with low success rate, such as pen and assembly, reflow may lead to worse performance. However, the SeFA-Policy retains high score when it comes to base policy 100% accuracy tasks, such as the plate-slide-side task. It proves our hypothesis that the performance degradation brought by reflow is a kind of error accumulation. Fortunately, our proposed SeFA technique compensates for accuracy loss and achieves a much higher score, sometimes even better than base policy. Note that π_0 [26] is equivalent to the first policy in Table IV. Although it is also trained via the flow matching loss, π_0 does not involve any reflow or *Selective Flow Alignment*, which effectiveness is underlined by the numbers in the last row in Table IV. This means that vanilla flow-based policies can be strengthened with our proposed methods, since ours can be implemented upon π_0 or any other policy that utilizes flow matching loss.

Ablation study on 1-step solver. A 1-step Euler solver works well in the default SeFA-Policy inference settings. To further explore it effectiveness, we replace it with a 100-step solver. Intuitively, a solver with more steps contributes to a higher accuracy since the truncation error introduced by using an approximation is decreased in every denoising loop. We also evaluate the performance of the Runge-Kutta method of order 5(4) from Scipy [43], denoted as RK45. The number of steps is adaptively decided based on user-specified relative and absolute tolerances, with its minimum no less than 1. As shown in Table V, the RK45 solver reports the highest accuracy for reflow policy. For SeFA-Policy, the 1-step solver is on a par with the 100-step one. Notwithstanding the comparable success rates, a 1-step solver exceeds its 100-step counterpart in inference efficiency by approximately two orders of magnitude. Since the number of sampling steps of RK45 is not fixed to 1, it still lags behind SeFA-Policy with respect to efficiency.

C. Efficiency

We evaluate inference latency on an NVIDIA RTX 6000 Ada GPU on the Meta-World assembly task. SeFA-Policy requires only a single denoising step per action prediction, while Diffusion Policy requires 100 steps as per original settings and AdaFlow 20. With observation encoding consuming negligible time compared to sampling, SeFA-Policy achieves dramatic speedup.

After a 200-iteration warm-up, we measure average latency over 800 rollouts, excluding the upper and lower duration quartiles. Table III presents wall clock times for each algorithm. SeFA-Policy achieves **98.7%** and **97.3%** acceleration compared to Diffusion Policy and AdaFlow Policy respectively, while maintaining superior accuracy across all domains. Consistency Policy achieves comparable inference speed with its 1-step sampling approach but requires extra training computation when distilling its teacher model. Consistency Policy underperforms compared to SeFA-Policy despite similar inference speed. 1-step DDIM inference shows comparable speed to SeFA-Policy but significantly sacrifices performance, failing in nearly all tasks.

D. Robustness

During evaluation, we observed that SeFA-Policy shows more robustness than Diffusion Policy and AdaFlow Policy. Take Adroit door and pen as examples. Evaluated on one checkpoint iteratively using different inference seeds, our algorithm sees a lower variance in accuracy than that of baselines. This is an extremely important attribute since in diffusion-based algorithms, different noise latent initializations are heavily related to the final performance. A well-generalized algorithm should maintain a steady success rate among different initial noise samples to avoid fluctuation in performance during random evaluation. The robustness of SeFA-Policy may be mainly attributed to the straight and stable sampling trajectories of rectified flow modeling. Results are shown in Figure 4.

V. Real World Experiments

To comprehensively evaluate our method in the real world, we select 7 representative manipulation tasks, as shown in Table VI. The tasks cover a variety of motions, including twist, pick, place, push, pull, and sweep, posing challenges in terms of precision, dexterity, and long-horizon.

TABLE I: **Main simulation results.** Averaged over 66 tasks, SeFA achieves **50.1**% relative improvement compared to Diffusion Policy.

Algorithm \ Task	Adroit (3)	RoboMimic (5)	Kitchen (7)	Push-T (1)	Meta-World Easy (28)	Meta-World Medium (11)	Meta-World Hard (6)	Meta-World Very Hard (5)	Average (66)
SeFA	**41.3**	**95.4**	**58.1**	**80.0**	**78.1**	**34.5**	**42.2**	**41.4**	**62.3 (↑ 50.1%)**
Diffusion Policy	25.0	93.2	57.1	78.0	38.3	19.5	16.7	32.0	38.9
AdaFlow Policy	33.7	95.0	49.2	72.0	46.2	26.7	17.8	12.0	41.5

TABLE II: **Comparing SeFA with more baselines in simulation.**

Algorithm \ Task	**Adroit** Pen	**Push-T**	**RoboMimic** Square	**Average**
SeFA	**52±8**	**80±2**	**97±5**	**76.3**
Diffusion Policy (100-step)	20±7	78±1	83±2	60.3
Diffusion Policy (1-step)	0±0	70±0	0±0	23.3
Consistency Policy (1-step)	32±8	71±2	89±2	64.0
AdaFlow Policy (20-step)	37±9	72±0	90±3	66.3

TABLE III: **Inference latency for one action prediction step in simulation (ms).**

Algorithm	NFE	Inference Latency (ms)
SeFA	1	16.72
Diffusion Policy	100	1287
Consistency Policy	1	18.35
AdaFlow Policy	20	629.0
1-step DDIM	1	16.16

TABLE IV: **Ablation on the selective flow alignment.**

Designs \ Task	#Steps	**Adroit** Pen	**Meta-World** Assembly	Plate Slide Side	**Average**
Base Policy	100	43±6	27±9	**100±0**	56.7
Reflow Policy [3]	1	40±8	7±9	**100±0**	49.0
SeFA	1	**52±8**	**33±9**	**100±0**	**61.7**

TABLE V: **Ablation on solver choices.** We compare the effectiveness of the 1-step Euler solver with a 100-step solver and a RK45 solver with adaptive timesteps.

Task\ Solver		**Adroit** Pen	**Meta-World** Assembly	PlateSlide	**Ave.**
Base Policy	1-step solver	43±6	27±9	**100±0**	57±5
	100-step solver	45±7	**30±6**	**100±0**	58±4
	RK45 solver	**46±7**	30±9	**100±0**	**59±5**
Reflow Policy	1-step solver	40±8	7±9	**100±0**	49±6
	100-step solver	40±4	25±9	**100±0**	55±4
	RK45 solver	**45±6**	**30±8**	**100±0**	**58±5**
SeFA-Policy	1-step solver	52±8	33±9	**100±0**	62±6
	100-step solver	**52±6**	**33±6**	**100±0**	**62±4**
	RK45 solver	50±4	31±4	**100±0**	60±3

TABLE VI: **Main results for real robot experiments.** The first 3 and last 4 tasks are evaluated with 10 and 20 trials each, respectively.

Real Robot Benchmark (7 Tasks)		
Task	Diffusion Policy	SeFA
Putting Apple in the Bowl	**100%**	**100%**
Moving Object Picking	60%	**70%**
Flower Insertion	20%	**80%**
Coffee Bean Sweeping	35%	**70%**
Drawer Close	40%	**60%**
Rice Pouring	0%	**30%**
Knob Pull	0%	**40%**

A. Experimental Analysis

Real-world manipulation results. Table VI presents a comparative analysis of success rates between SeFA-Policy and Diffusion Policy across various real-world robotic manipulation tasks. The empirical evidence demonstrates that SeFA-Policy consistently outperforms Diffusion Policy in tasks demanding high precision, dexterity, and long-horizon sequential manipulation. Diffusion Policy exhibits inferior performance primarily because it is inherently unstable, with success rates substantially varying across different noise initializations. In the Putting Apples in the Bowl task, both policies achieve perfect success rates (100%), indicating comparable capabilities in handling basic pick-and-place operations with irregularly shaped objects. However, the performance gap widens significantly in favor of SeFA-Policy as task complexity increases. For the Moving Object Picking, which involves handling a rubber duck on a dynamic water surface, SeFA-Policy achieves a 70% success rate compared to Diffusion Policy's 60%. This superiority can be attributed to the continuous rectification mechanism in SeFA-Policy, which facilitates more robust and adaptive visuomotor control when managing complex interactions between the gripper, floating objects, and water surface. In the most challenging tasks—Flower Insertion and Rice Pouring—which involve long-horizon sequential manipulation consisting of precise grasping and delicate wrist twisting, SeFA-Policy achieves 80% and 30% success rate, substantially outperforming Diffusion Policy's 20% and 10%. This marked improvement underscores the efficacy of SeFA-Policy in gen-

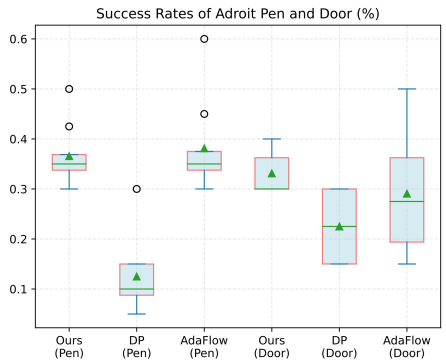
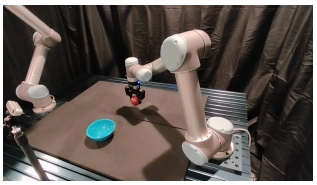
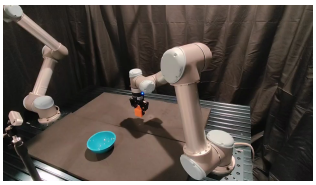


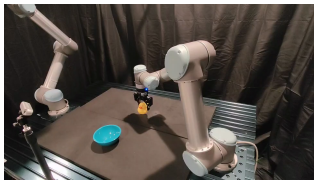
Fig. 4: **Success Rates on Adroit (%).**



(a) **Apple.**



(b) **Orange cube.**



(c) **Rubber duck.**

Fig. 5: **Grasping different objects with one policy.** SeFA trained on the apple can generalize to other objects (cube, rubber duck) with similar sizes and locations.

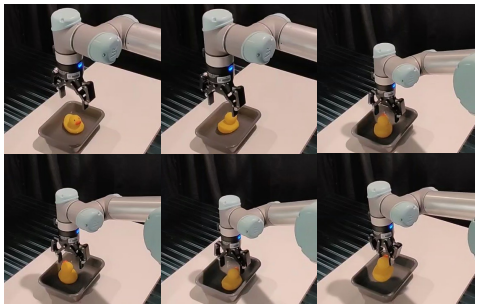


Fig. 6: **Floating object manipulation.** SeFA dynamically adjusts its action trajectory to approach and grab the moving rubber duck on the water, which demonstrates generalization ability to different object locations.

erating consistent and accurate action sequences for multi-stage, fine-grained continuous manipulation scenarios.

Handling different object appearance. As shown in Figure 5, while trained on grasping one object, the policy successfully generalizes to others in various appearances, suggesting that SeFA-Policy can leverage learned visuomotor patterns to handle variations in object appearance while maintaining successful grasp execution.

Manipulating dynamic objects. Figure 6 illustrates how SeFA-Policy adjusts its action trajectory to approach and grasp a moving rubber duck on water. Unlike static object grasping, this task requires the policy to continuously refine its motion based on the ever-changing real-time position. The sequential images indicate that SeFA-Policy can generate high-frequency visuomotor actions to compensate for object drift, allowing successful execution even towards unfixed location. This highlights its suitability for real-world tasks involving dynamic objects.

Precise manipulation. Flower Insertion and Rice Pouring involve multi-stage manipulation, i.e., first grasping and then performing twists in one or several joints. Rice Pouring is more demanding, which requires an exact wrist twisting whilst other joints remain still to ensure the rice grains fall within the bowl. Knob Pull and Drawer Close require the end effector to perform a pull/push towards a straight line, otherwise the drawer would stuck by friction. Also, Knob Pull includes precise grasping before the smooth pull action. In Coffee Bean Sweeping, the robot arm should apply a force onto the tabletop, which needs precise height control.

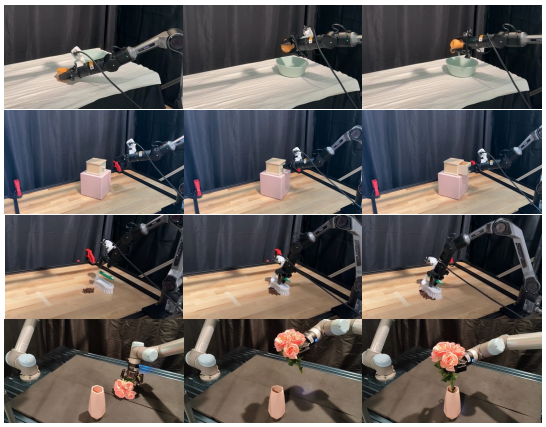


Fig. 7: **Precise manipulation.** SeFA successfully performs Rice Pouring, Knob Pull, Coffee Bean Sweeping, and Flower Insertion (from top to bottom).

VI. Conclusion

In this work, we presented SeFA, visuomotor policy learning framework that supports efficient and accurate inference. SeFA is a flow-based model that facilitates single-step action synthesis and addresses the performance degradation by introducing a novel selective flow alignment strategy. Through rigorous real-world and simulation evaluations, SeFA exhibited remarkable performance across tasks demanding high precision, dexterity, and extended temporal planning horizons. Comparative analyses demonstrate that SeFA substantially outperforms state-of-the-art approaches, achieving superior success rates while concurrently reducing inference latency by 98.7%, thereby enabling significantly more responsive and stable robot execution. These empirical findings underscore the efficacy of SeFA as a computationally efficient and highly effective solution for real-time visuomotor control in sophisticated manipulation tasks, with promising implications for future deployments in dynamic robotic applications.

References

[1] C. Chi, S. Feng, Y. Du, Z. Xu, E. Cousineau, B. Burchfiel, and S. Song, "Diffusion policy: Visuomotor policy learning via action diffusion," *arXiv preprint arXiv:2303.04137*, 2023.

[2] X. Hu, B. Liu, X. Liu, and Q. Liu, "Adaflow: Imitation learning with variance-adaptive flow-based policies," *arXiv preprint arXiv:2402.04292*, 2024.

[3] X. Liu, C. Gong, and qiang liu, "Flow straight and fast: Learning to generate and transfer data with rectified flow," in *The Eleventh International Conference on Learning Representations*, 2023. [Online]. Available: https://openreview.net/forum?id=XVjTT1nw5z

[4] M. Janner, Y. Du, J. Tenenbaum, and S. Levine, "Planning with diffusion for flexible behavior synthesis," in *Proceedings of the 39th International Conference on Machine Learning*. PMLR, 2022, pp. 9902–9915, ISSN: 2640-3498. [Online]. Available: https://proceedings.mlr.press/v162/janner22a.html

[5] Y. Luo, C. Sun, J. B. Tenenbaum, and Y. Du, "Potential based diffusion motion planning," in *Forty-first International Conference on Machine Learning*, 2024.

[6] J. Carvalho, A. Le, M. Baierl, D. Koert, and J. Peters, "Motion planning diffusion: Learning and planning of robot motions with diffusion models," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023.

[7] K. Saha, V. Mandadi, J. Reddy, A. Srikanth, A. Agarwal, B. Sen, A. Singh, and M. Krishna, "Edmp: Ensemble-of-costs-guided diffusion for motion planning," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 10 351–10 358.

[8] S. Huang, Z. Wang, P. Li, B. Jia, T. Liu, Y. Zhu, W. Liang, and S.-C. Zhu, "Diffusion-based generation, optimization, and planning in 3d scenes," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

[9] T. Pearce, T. Rashid, A. Kanervisto, D. Bignell, M. Sun, R. Georgescu, S. V. Macua, S. Z. Tan, I. Momennejad, K. Hofmann, and S. Devlin, "Imitating human behaviour with diffusion models," in *The Eleventh International Conference on Learning Representations*, 2023.

[10] C. Chi, S. Feng, Y. Du, Z. Xu, E. Cousineau, B. Burchfiel, and S. Song, "Diffusion policy: Visuomotor policy learning via action diffusion," in *Robotics: Science and Systems 2023*. Robotics: Science and Systems Foundation. [Online]. Available: http://www.roboticsproceedings.org/rss19/p026.pdf

[11] H. Ha, P. Florence, and S. Song, "Scaling up and distilling down: Language-guided robot skill acquisition," in *Proceedings of The 7th Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, J. Tan, M. Toussaint, and K. Darvish, Eds., vol. 229. PMLR, 06–09 Nov 2023, pp. 3766–3777. [Online]. Available: https://proceedings.mlr.press/v229/ha23a.html

[12] Z. Xian, N. Gkanatsios, T. Gervet, T.-W. Ke, and K. Fragkiadaki, "Chaineddiffuser: Unifying trajectory diffusion and keypose prediction for robotic manipulation," in *Proceedings of The 7th Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, J. Tan, M. Toussaint, and K. Darvish, Eds., vol. 229. PMLR, 06–09 Nov 2023, pp. 2323–2339. [Online]. Available: https://proceedings.mlr.press/v229/xian23a.html

[13] X. Li, V. Belagali, J. Shang, and M. S. Ryoo, "Crossway diffusion: Improving diffusion-based visuomotor policy via self-supervised learning." [Online]. Available: http://arxiv.org/abs/2307.01849

[14] Y. Ze, G. Zhang, K. Zhang, C. Hu, M. Wang, and H. Xu, "3d diffusion policy: Generalizable visuomotor policy learning via simple 3d representations," in *Proceedings of Robotics: Science and Systems (RSS)*, 2024.

[15] L. Wang, J. Zhao, Y. Du, E. Adelson, and R. Tedrake, "PoCo: Policy Composition from and for Heterogeneous Robot Learning," in *Proceedings of Robotics: Science and Systems*, Delft, Netherlands, July 2024.

[16] K. Chen, E. Lim, L. Kelvin, Y. Chen, and H. Soh, "Don't Start From Scratch: Behavioral Refinement via Interpolant-based Policy Diffusion," in *Proceedings of Robotics: Science and Systems*, Delft, Netherlands, July 2024.

[17] K. Sridhar, S. Dutta, D. Jayaraman, J. Weimer, and I. Lee, "Memory-consistent neural networks for imitation learning," 2023.

[18] T. Z. Zhao, J. Tompson, D. Driess, P. Florence, S. K. S. Ghasemipour, C. Finn, and A. Wahid, "ALOHA unleashed: A simple recipe for robot dexterity," in *8th Annual Conference on Robot Learning*, 2024. [Online]. Available: https://openreview.net/forum?id=gvdXE7ikHI

[19] C. Chi, Z. Xu, C. Pan, E. Cousineau, B. Burchfiel, S. Feng, R. Tedrake, and S. Song, "Universal Manipulation Interface: In-The-Wild Robot Teaching Without In-The-Wild Robots," in *Proceedings of Robotics: Science and Systems*, Delft, Netherlands, July 2024.

[20] M. Reuss, M. Li, X. Jia, and R. Lioutikov, "Goal-conditioned imitation learning using score-based diffusion policies," in *Robotics: Science and Systems 2023*. Robotics: Science and Systems Foundation. [Online]. Available: http://www.roboticsproceedings.org/rss19/p028.pdf

[21] M. Reuss, Ömer Erdinç Yağmurlu, F. Wenzel, and R. Lioutikov, "Multimodal Diffusion Transformer: Learning Versatile Behavior from Multimodal Goals," in *Proceedings of Robotics: Science and Systems*, Delft, Netherlands, July 2024.

[22] L. Chen, S. Bahl, and D. Pathak, "Playfusion: Skill acquisition via diffusion from language-annotated play," in *CoRL*, 2023.

[23] E. Zhang, Y. Lu, W. Wang, and A. Zhang, "Lad: Language control diffusion: efficiently scaling through space, time, and tasks," *arXiv preprint arXiv:2210.15629*, 2023.

[24] J. Urain, N. Funk, J. Peters, and G. Chalvatzaki, "Se(3)-diffusionfields: Learning smooth cost functions for joint grasp and motion optimization through diffusion," *IEEE International Conference on Robotics and Automation (ICRA)*, 2023.

[25] Y. Lipman, R. T. Chen, H. Ben-Hamu, M. Nickel, and M. Le, "Flow matching for generative modeling," in *The Eleventh International Conference on Learning Representations*, 2022.

[26] K. Black, N. Brown, D. Driess, A. Esmail, M. Equi, C. Finn, N. Fusai, L. Groom, K. Hausman, B. Ichter *et al.*, "pi_0: A vision-language-action flow model for general robot control," *arXiv preprint arXiv:2410.24164*, 2024.

[27] T. Karras, M. Aittala, T. Aila, and S. Laine, "Elucidating the design space of diffusion-based generative models," vol. 35, pp. 26 565–26 577.

[28] Y. Song, P. Dhariwal, M. Chen, and I. Sutskever, "Consistency models," in *Proceedings of the 40th International Conference on Machine Learning*. PMLR, pp. 32 211–32 252, ISSN: 2640-3498. [Online]. Available: https://proceedings.mlr.press/v202/song23a.html

[29] J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models." [Online]. Available: https://openreview.net/forum?id=St1giarCHLP

[30] D. Kim, C.-H. Lai, W.-H. Liao, N. Murata, Y. Takida, T. Uesaka, Y. He, Y. Mitsufuji, and S. Ermon, "Consistency trajectory models: Learning probability flow ODE trajectory of diffusion," in *The Twelfth International Conference on Learning Representations*, 2024. [Online]. Available: https://openreview.net/forum?id=ymjI8feDTD

[31] P. M. Scheikl, N. Schreiber, C. Haas, N. Freymuth, G. Neumann, R. Lioutikov, and F. Mathis-Ullrich, "Movement primitive diffusion: Learning gentle robotic manipulation of deformable objects," *IEEE Robotics and Automation Letters*, vol. 9, no. 6, pp. 5338–5345, 2024.

[32] S. H. Høeg, Y. Du, and O. Egeland, "Streaming diffusion policy: Fast policy synthesis with variable noise diffusion models," 2024. [Online]. Available: https://arxiv.org/abs/2406.04806

[33] D. Ruhe, J. Heek, T. Salimans, and E. Hoogeboom, "Rolling diffusion models," in *Proceedings of the 41st International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, R. Salakhutdinov, Z. Kolter, K. Heller, A. Weller, N. Oliver, J. Scarlett, and F. Berkenkamp, Eds., vol. 235. PMLR, 21–27 Jul 2024, pp. 42 818–42 835. [Online]. Available: https://proceedings.mlr.press/v235/ruhe24a.html

[34] Z. Zhang, R. Liu, R. Hanocka, and K. Aberman, "Tedi: Temporally-entangled diffusion for long-term motion synthesis," in *ACM SIGGRAPH 2024 Conference Papers*, ser. SIGGRAPH '24. New York, NY, USA: Association for Computing Machinery, 2024. [Online]. Available: https://doi.org/10.1145/3641519.3657515

[35] Y. Song, J. Sohl-dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," 2021. [Online]. Available: https://openreview.net/forum?id=PxTIG12RRHS

[36] A. Rajeswaran, V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, and S. Levine, "Learning complex dexterous manipulation with deep reinforcement learning and demonstrations," *arXiv*, 2017.

[37] A. Mandlekar, D. Xu, J. Wong, S. Nasiriany, C. Wang, R. Kulkarni, L. Fei-Fei, S. Savarese, Y. Zhu, and R. Martín-Martín, "What matters in learning from offline human demonstrations for robot manipulation," in *5th Annual Conference on Robot Learning*, 2021.

[38] T. Yu, D. Quillen, Z. He, R. Julian, K. Hausman, C. Finn, and S. Levine, "Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning," in *CoRL*, 2020.

[39] A. Gupta, V. Kumar, C. Lynch, S. Levine, and K. Hausman, "Relay policy learning: Solving long horizon tasks via imitation and reinforcement learning," *Conference on Robot Learning (CoRL)*, 2019.

[40] P. Florence, C. Lynch, A. Zeng, O. A. Ramirez, A. Wahid, L. Downs, A. Wong, J. Lee, I. Mordatch, and J. Tompson, "Implicit behavioral cloning," in *CoRL*, 2022.

[41] C. Chi, S. Feng, Y. Du, Z. Xu, E. Cousineau, B. C. Burchfiel, and S. Song, "Diffusion Policy: Visuomotor Policy Learning via Action Diffusion," in *Proceedings of Robotics: Science and Systems*, Daegu, Republic of Korea, July 2023.

[42] A. Prasad, K. Lin, J. Wu, L. Zhou, and J. Bohg, "Consistency policy: Accelerated visuomotor policies via consistency distillation," 2024. [Online]. Available: http://arxiv.org/abs/2405.07503

[43] "Scipy rk45 function," https://docs.scipy.org/doc/scipy/reference/generated/scipy.integrate.RK45.html, accessed: 2022-08-19.