

# Novelty Assessment Report

**Paper:** Mean Flow Policy with Instantaneous Velocity Constraint for One-step Action Generation

**PDF URL:** <https://openreview.net/pdf?id=mleKe74W43>

**Venue:** ICLR 2026 Conference Submission

**Year:** 2026

**Report Generated:** 2025-12-29

## Abstract

Learning expressive and efficient policy functions is a promising direction in reinforcement learning (RL). While flow-based policies have recently proven effective in modeling complex action distributions with a fast deterministic sampling process, they still face a trade-off between expressiveness and computational burden, which is typically controlled by the number of flow steps. In this work, we propose mean flow policy (MFP), a new generative policy function that models the mean velocity field to achieve the fastest one-step action generation. To ensure its high expressiveness, an instantaneous velocity constraint (IVC) is introduced on the mean velocity field during training. We theoretically prove that this design explicitly serves as a crucial boundary condition, thereby improving learning accuracy and enhancing policy expressiveness. Empirically, our MFP achieves state-of-the-art success rates across several challenging robotic manipulation tasks from Robomimic and OGBench. It also delivers substantial improvements in training and inference speed over existing flow-based policy baselines.

### Disclaimer

This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.

Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.

If you have any questions, please contact: mingzhang23@m.fudan.edu.cn

## Core Task Landscape

This paper addresses: **One-Step Action Generation for Robotic Manipulation Using Flow-Based Policies**

A total of **28 papers** were analyzed and organized into a taxonomy with **15 categories**.

### Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Flow Matching and Velocity Field Learning**
- **Acceleration and Distillation Techniques**
- **Multi-Step Flow Inference and Regularization**
- **Reinforcement Learning Integration**
- **Application-Specific Flow Policies**

### Complete Taxonomy Tree

- One-Step Action Generation for Robotic Manipulation Using Flow-Based Policies Survey Taxonomy
- Flow Matching and Velocity Field Learning
  - Mean Velocity Field Approaches ★ (2 papers)
  - [0] Mean Flow Policy with Instantaneous Velocity Constraint for One-step Action Generation (Anon et al., 2026) [View paper](#)
  - [4] MP1: MeanFlow Tames Policy Learning in 1-step for Robotic Manipulation (Wang, 2025) [View paper](#)
  - Consistency Flow Matching (2 papers)
  - [3] Flowpolicy: Enabling fast and robust 3d flow-based policy via consistency flow matching for robot manipulation (Qinglun Zhang, 2025) [View paper](#)
  - [6] Manifold: A general robot manipulation policy via consistency flow training (Yan Ge, 2025) [View paper](#)
  - Rectified Flow Models (2 papers)
  - [27] Fast and Accurate Visuomotor Imitation Learning via 2D Consistency Flow Matching Policy (S Tian, 2025) [View paper](#)
  - [28] RecFlow Policy: Fast and Accurate Visuomotor Policy Learning via Rectified Action Flow (R Xue, n.d.) [View paper](#)
  - Residual Flow Reformulations (1 papers)
  - [21] One-Step Generative Policies with Q-Learning: A Reformulation of MeanFlow (Zeyuan Wang, 2025) [View paper](#)
- Acceleration and Distillation Techniques
  - Variance-Adaptive and Completion-Based Acceleration (2 papers)
  - [1] Adaflow: Imitation learning with variance-adaptive flow-based policies (Hu Xixi, 2024) [View paper](#)
  - [5] Flow-Based Single-Step Completion for Efficient and Expressive Policy Learning (Fleming Cody, 2025) [View paper](#)
  - Frequency and Alignment-Based Acceleration (2 papers)
  - [10] FreqPolicy: Efficient Flow-based Visuomotor Policy via Frequency Consistency (Su Yifei, 2025) [View paper](#)
  - [24] SeFA-Policy: Fast and Accurate Visuomotor Policy Learning with Selective Flow Alignment (Rong Xue, 2025) [View paper](#)
  - Optimal Transport and Sample Flow Methods (2 papers)
  - [25] L1 Sample Flow for Efficient Visuomotor Learning (Weixi Song, 2025) [View paper](#)
  - [26] Fast Flow-based Visuomotor Policies via Conditional Optimal Transport Couplings (Malkin, 2025) [View paper](#)
- Multi-Step Flow Inference and Regularization
  - Dispersive Regularization and Representation Collapse (1 papers)
  - [12] DM1: MeanFlow with Dispersive Regularization for 1-Step Robotic Manipulation (Zou Guowei, 2025) [View paper](#)
  - Non-Uniform Scheduling and Streaming Flow (2 papers)
  - [17] Dense-Jump Flow Matching with Non-Uniform Time Scheduling for Robotic Policies: Mitigating Multi-Step Inference Degradation (Chen Zi-dong, 2025) [View paper](#)

- [19] Streaming Flow Policy: Simplifying diffusion/flow-matching policies by treating action trajectories as flow trajectories (Fang Xiao-lin, 2025) [View paper](#)
- Reinforcement Learning Integration
  - Online RL Fine-Tuning and Mirror Descent (2 papers)
  - [13] ReinFlow: Fine-tuning flow matching policy with online reinforcement learning (Zhang Tonghe, 2025) [View paper](#)
  - [20] One-step flow policy mirror descent (Chen, 2025) [View paper](#)
  - Q-Learning Integration and Policy Composition (3 papers)
  - [18] Revisiting Diffusion Q-Learning: From Iterative Denoising to One-Step Action Generation (Nguyen, 2025) [View paper](#)
  - [22] Decision Flow Policy Optimization (HU Jifeng, 2025) [View paper](#)
  - [23] Compose Your Policies! Improving Diffusion-based or Flow-based Robot Policies via Test-time Distribution-level Composition (Cao, 2025) [View paper](#)
- Application-Specific Flow Policies
  - 3D Perception and Region-Aware Policies (2 papers)
  - [7] FlowRAM: Grounding Flow Matching Policy with Region-Aware Mamba Framework for Robotic Manipulation (Sen Wang, 2025) [View paper](#)
  - [8] 3D FlowMatch Actor: Unified 3D Policy for Single- and Dual-Arm Manipulation (Gkanatsios Nikolaos, 2025) [View paper](#)
  - Bimanual and Cloth Manipulation Policies (2 papers)
  - [14] Towards a Generalizable Bimanual Foundation Policy via Flow-based Video Prediction (Fan, 2025) [View paper](#)
  - [15] FabricFlowNet: Bimanual Cloth Manipulation with a Flow-based Policy (Thomas Weng, 2021) [View paper](#)
  - Multimodal and Cross-Domain Policies (3 papers)
  - [2] Flow as the cross-domain manipulation interface (XU Mengda, 2024) [View paper](#)
  - [9] FlowDreamer: A RGB-D World Model with Flow-based Motion Representations for Robot Manipulation (Guo Jun, 2025) [View paper](#)
  - [11] FBI: Learning Dexterous In-hand Manipulation with Dynamic Visuotactile Shortcut Policy (Chen Yijin, 2025) [View paper](#)
  - Motion Primitives and Temporal Consistency (1 papers)
  - [16] FRMD: Fast Robot Motion Diffusion with Consistency-Distilled Movement Primitives for Smooth Action Generation (Jin Jun, 2025) [View paper](#)

## Narrative

Core task: one-step action generation for robotic manipulation using flow-based policies. The field has coalesced around learning continuous normalizing flows that map noise distributions to action distributions, enabling fast inference without iterative denoising. The taxonomy reveals several complementary research directions. Flow Matching and Velocity Field Learning focuses on foundational training objectives—such as learning mean velocity fields or consistency-based formulations—that define how the flow evolves from a base distribution to the target policy. Acceleration and Distillation Techniques address inference speed by distilling multi-step flows into fewer steps or even single-step generators, often trading off some expressiveness for real-time performance. Multi-Step Flow Inference and Regularization explores structured sampling strategies and constraints that improve stability or sample quality when multiple flow steps are retained. Reinforcement Learning Integration examines how flow-based policies can be optimized via RL objectives, blending generative modeling with value-based or policy-gradient methods. Finally, Application-Specific Flow Policies tailors flow architectures to particular domains such as dexterous manipulation, bimanual coordination, or fabric handling.

Recent work has intensified efforts to achieve true one-step generation while preserving multimodal expressiveness. Mean Flow Velocity[0] and MP1 MeanFlow[4] exemplify mean velocity field approaches that directly predict the average flow direction, enabling efficient single-step inference without sacrificing the ability to capture diverse action modes. In contrast, Flowpolicy Consistency[3] and Flow Single Step[5] pursue consistency distillation or rectified flow training to collapse multi-step trajectories into a single forward pass, often at the cost of additional training complexity or slight mode collapse risks. Meanwhile, methods like Manifold[6] and FlowRAM[7] incorporate manifold constraints or memory-augmented architectures to handle high-dimensional or temporally extended action spaces. Mean Flow Velocity[0] sits squarely within the mean velocity field cluster, sharing the goal of direct one-step prediction with MP1 MeanFlow[4] but differing in how the mean field is regularized or conditioned on observations, offering a streamlined alternative to the iterative refinement seen in consistency-based approaches like Flowpolicy Consistency[3].

## Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

### 1. MP1: MeanFlow Tames Policy Learning in 1-step for Robotic Manipulation

**Authors:** Wang, Ziyi, Juyi Sheng, Li Peiming, Ziyi Wang, et al. (9 authors total) | **Year/Venue:** 2025 | **URL:** [View paper](#)

#### Abstract

In robot manipulation, robot learning has become a prevailing approach. However, generative models within this field face a fundamental trade-off between the slow, iterative sampling of diffusion models and the architectural constraints of faster Flow-based methods, which often rely on explicit consistency losses. To address these limitations, we introduce MP1, which pairs 3D point-cloud inputs with the MeanFlow paradigm to generate action trajectories in one network function evaluation (1-NFE)...

#### Relationship Analysis

Both papers belong to the Mean Velocity Field Approaches category, learning interval-averaged velocity fields for one-step action generation in robotic manipulation. They share the core MeanFlow paradigm that models mean velocity over time intervals to achieve 1-NFE inference, eliminating multi-step sampling overhead. The key difference is that the original paper (MFP) focuses on theoretical boundary conditions via instantaneous velocity constraint (IVC) for offline-to-online RL with Q-learning, while the candidate paper (MP1) emphasizes 3D point-cloud conditioning with a Dispersive Loss for few-shot generalization in imitation learning settings.

## Contributions Analysis

**Overall novelty summary.** The paper proposes a mean flow policy (MFP) that learns a mean velocity field for one-step action generation in robotic manipulation. It resides in the 'Mean Velocity Field Approaches' leaf, which contains only two papers including this one. This is a relatively sparse research direction within the broader taxonomy of 28 papers across 15 leaf nodes, suggesting the specific formulation of mean velocity fields for direct one-step generation remains an emerging area compared to more crowded branches like consistency flow matching or acceleration techniques.

The taxonomy reveals several neighboring directions. Consistency Flow Matching (4 papers) and Rectified Flow Models (2 papers) pursue similar one-step or few-step generation goals but through different training objectives—consistency distillation or straight-trajectory rectification rather than mean velocity field learning. The Acceleration and Distillation branch (6 papers) addresses speed through post-hoc distillation or variance adaptation, while this work aims for inherent one-step efficiency. The scope notes clarify that

mean velocity field methods exclude consistency-based reformulations, positioning MFP as a distinct approach within the flow matching paradigm.

Among 30 candidates examined, the core contribution of mean flow policy for one-step generation shows substantial prior work overlap: 5 of 10 candidates examined appear refutable, indicating existing methods pursue similar one-step generation goals. The instantaneous velocity constraint (IVC) as a boundary condition shows no clear refutation across 10 candidates, suggesting this theoretical contribution may be more novel. The performance claim examined 10 candidates with 2 refutable, implying competitive baselines exist but the specific speedup-accuracy trade-off may differ. The limited search scope means these statistics reflect top-30 semantic matches, not exhaustive coverage.

Based on the top-30 candidate analysis, the work appears to offer incremental refinement within an emerging subfield. The mean velocity field formulation has at least one sibling paper and multiple related approaches in neighboring leaves, while the IVC boundary condition lacks clear precedent in the examined literature. The sparse taxonomy leaf suggests room for methodological diversity, though the refutation statistics indicate the core one-step generation goal is not unprecedented.

---

This paper presents **3 main contributions**, each analyzed against relevant prior work:

### Contribution 1: Mean Flow Policy for one-step action generation

**Description:** The authors introduce a novel flow-based policy that models the mean velocity field instead of instantaneous velocities. This design enables direct single-step mapping from Gaussian noise to multi-modal action distributions, eliminating the multi-step iterative sampling overhead of existing flow policies while preserving expressive power.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

#### 1. MP1: MeanFlow Tames Policy Learning in 1-step for Robotic Manipulation

URL: [View paper](#)

##### Prior Art Analysis

MP1 MeanFlow[4] demonstrates that the mean velocity field approach for one-step action generation was already applied to robotic manipulation before the original paper. Both papers model the mean velocity field  $u(a(t), t, r, s)$  using the same mathematical formulation and derive it from the mean flow identity. The candidate paper explicitly states it is 'the first adaptation of the meanflow paradigm to robot learning' and provides the identical mathematical framework, including the mean velocity definition, mean flow identity, and one-step inference formula that the original paper claims as novel.

##### Evidence

Evidence 1 - **Rationale:** The candidate explicitly claims to be the first to adapt meanflow to robot learning, directly challenging the original's novelty claim of proposing a 'new generative policy function' based on mean velocity fields. - **Original:** we propose mean flow policy (mfp), a new generative policy function that models the mean velocity field to achieve the fastest one-step action generation - **Candidate:** we introduce mp1, the first meanflow-based robot-learning framework. conditioning on 3d point-cloud features, it learns effectively from a handful of demonstrations, yet delivers one-step sampling with sota success rates and millisecond-level inference latency

Evidence 2 - **Rationale:** Both papers describe the same core innovation: learning mean/average velocity fields instead of instantaneous velocities to enable one-step generation, using nearly identical technical descriptions. - **Original:** while existing flow policies learn instantaneous velocities and require multi-step iterative sampling (lipman et al., 2023; park et al., 2025; bharadhwaj et al., 2024), mfp instead learns the mean velocity field (geng et al., 2025a). this design enables a direct, single-step mapping from a base gaus... - **Candidate:** unlike standard flow matching (fm), which learns an instantaneous velocity field  $v(zt, t)$  and requires solving an ode for sampling, meanflow learns the average velocity field  $u(zt, r, t)$  over an interval  $[r, t]$ . the average velocity is defined as the total displacement divided by the time duration

Evidence 3 - **Rationale:** The mathematical definition of mean velocity is identical in both papers, showing the same technical approach was already published in the candidate. - **Original:**  $u(a(t), t, r, s) \triangleq \frac{1}{t-r} \int_r^t v(a(\tau), \tau, s) d\tau$  - **Candidate:**  $u(zt, r, t) \triangleq \frac{1}{t-r} \int_r^t v(z\tau, \tau, s) d\tau$

Evidence 4 - **Rationale:** Both papers derive and use the same mean flow identity equation, which is the core mathematical foundation for their approach. - **Original:**  $-u(a(t), t, r, s) + (r-t) \frac{d}{dt} u(a(t), t, r, s) = -v(a(t), t, s)$  - **Candidate:**  $u(zt, r, t) = v(zt, t) - (t-r) \frac{d}{dt} u(zt, r, t)$

Evidence 5 - **Rationale:** Both papers present the same one-step inference formula, showing the candidate already implemented this approach before the original paper's submission. - **Original:**  $a(1) = a(0) + u^*(a(0), 0, 1, s), a(0) \sim n(0, i)$  - **Candidate:**  $a_0 = a_1 - ucfg \theta(a_1, 0, 1, c)$

---

#### 2. Mean Flows for One-step Generative Modeling

URL: [View paper](#)

##### Prior Art Analysis

Mean Flows Generative[43] demonstrates that the concept of modeling mean velocity fields (termed 'average velocity' in their work) for one-step generation was already established prior to the original paper. Both papers introduce the same fundamental approach: replacing instantaneous velocity modeling with mean/average velocity modeling to enable direct single-step mapping from noise to target distributions. The candidate paper explicitly states they 'introduce the notion of average velocity to characterize flow fields, in contrast to instantaneous velocity modeled by flow matching methods' and achieve one-step generation without multi-step sampling. This directly challenges the novelty claim that the original authors were first to propose modeling mean velocity fields for one-step action generation, as the candidate demonstrates this concept was already developed in the generative modeling domain.

##### Evidence

Evidence 1 - **Rationale:** Both papers introduce the same core concept of modeling mean/average velocity instead of instantaneous velocity for one-step generation. The candidate explicitly introduces this notion and derives the identity between average and instantaneous velocities, which is the same fundamental approach used in the original paper's mean flow policy. - **Original:** we propose the mean flow policy (mfp) as an affirmative answer. while existing flow policies learn instantaneous velocities and require multi-step iterative sampling (lipman et al., 2023; park et al., 2025; bharadhwaj et al., 2024), mfp instead learns the mean velocity field (geng et al., 2025a). th... - **Candidate:** we introduce the notion of average velocity to characterize flow fields, in contrast to instantaneous velocity modeled by flow matching methods. a well-defined identity between average and instantaneous velocities is derived and used to guide neural network training. our method, termed the meanflow ...

Evidence 2 - **Rationale:** The original paper's mathematical formulation of mean velocity field is conceptually identical to the candidate's 'average velocity' approach. Both contrast their method against instantaneous velocity modeling used in standard flow matching. - **Original:** unlike standard flow policies that learn the instantaneous velocity field  $v(a(t), t, s)$ , we center on the mean velocity field  $u(a(t), t, r, s)$ , modeling the mean velocity over any given time interval  $[t, r]$  as  $u(a(t), t, r, s) \triangleq \frac{1}{t-r} \int_r^t v(a(\tau), \tau, s) d\tau$ . - **Candidate:** we introduce the notion of average velocity to characterize flow fields, in contrast to instantaneous velocity modeled by flow matching methods.

Evidence 3 - **Rationale:** Both papers demonstrate that their mean/average velocity approach enables single-step generation (1-nfe). The candidate's achievement of one-step generation through average velocity modeling predates the original paper's claim of being first to enable one-step action generation via mean velocity. - **Original:** if the mean velocity model  $u^*$  is ideally learned, the policy inference is formulated as  $a(1) = a(0) + u^*(a(0), 0, 1, s)$ ,  $a(0) \sim n(0, i)$ . - **Candidate:** meanflow demonstrates strong empirical performance: it achieves an fid of 3.43 with a single function evaluation (1-nfe) on imagenet 256x256 trained from scratch, significantly outperforming previous state-of-the-art one-step diffusion/flow models.

---

### 3. One-Step Generative Policies with Q-Learning: A Reformulation of MeanFlow

URL: [View paper](#)

#### Prior Art Analysis

One Step Q Learning[21] demonstrates that prior work exists on one-step generative policies modeling mean velocity fields for action generation. The candidate paper explicitly reformulates meanflow (Geng et al. 2025) for RL applications, introducing a residual formulation  $g(a, b, t) = a - u(a, b, t)$  that enables direct noise-to-action mapping. Both papers address the same core challenge: achieving one-step action generation while maintaining expressiveness through mean velocity field modeling, rather than instantaneous velocity fields used in traditional flow policies.

#### Evidence

Evidence 1 - **Rationale:** Both papers explicitly propose one-step generative policies based on meanflow reformulation, demonstrating that the candidate addresses the same core contribution. - **Original:** we propose mean flow policy (mfp), a new generative policy function that models the mean velocity field to achieve the fastest one-step action generation - **Candidate:** we propose a one-step generative framework for offline rl that enables expressive and stable policy learning without relying on multi-step integration or policy distillation. through a residual reformulation of meanflow, our method supports direct q-learning with multimodal policies via a single-sta...

Evidence 2 - **Rationale:** Both papers distinguish their approach from instantaneous velocity methods by modeling mean/average velocity fields, citing the same foundational work (Geng et al. 2025). - **Original:** while existing flow policies learn instantaneous velocities and require multi-step iterative sampling (lipman et al., 2023; park et al., 2025; bharadhwaj et al., 2024), mfp instead learns the mean velocity field (geng et al., 2025a). this design enables a direct, single-step mapping from a base gaus... - **Candidate:** meanflow.meanflow (geng et al. 2025) builds on flow matching by introducing an average velocity field  $u(a, b, t)$ . for any  $0 \leq b < t \leq 1$ , the average velocity is defined as the time-average of the instantaneous velocity  $v$  over the interval  $[b, t]$ :  $u(a, b, t) = \int_b^t v(a, \tau) d\tau$

Evidence 3 - **Rationale:** Both papers describe one-step inference mechanisms using learned mean velocity models, eliminating the need for iterative integration. - **Original:** if the mean velocity model  $u^*$  is ideally learned, the policy inference is formulated as  $a(1) = a(0) + u^*(a(0), 0, 1, s)$ ,  $a(0) \sim n(0, i)$  - **Candidate:** this formula generates  $\hat{a}$  via one-step velocity estimation using the learned  $u$ , without integrating over time. flow policies in rl.recent works have proposed using flow matching to learn policies in rl. in suchflow policies, the policy  $\pi_\theta$  is implicitly defined by a state-conditional velocity field  $v$ ...

---

### 4. Revisiting Diffusion Q-Learning: From Iterative Denoising to One-Step Action Generation

URL: [View paper](#)

#### Prior Art Analysis

Diffusion Q Learning[18] demonstrates that the concept of modeling average/mean velocity fields for one-step action generation was already established in prior work. The candidate paper explicitly models an 'average velocity field' using the same mathematical formulation (equation 10) and mean flow identity (equation 11) as the original paper. Both papers cite and build upon the same foundational work (Geng et al., 2025) for mean flow modeling. The candidate paper's approach to learning average velocity over arbitrary time intervals  $[r, t]$  and using the mean flow identity for training directly parallels the original paper's methodology, indicating that this technical approach was not novel to the original authors.

#### Evidence

Evidence 1 - **Rationale:** Both papers explicitly model mean/average velocity fields for one-step generation, with the candidate paper providing the same mathematical definition and citing the same foundational work (Geng et al., 2025). - **Original:** While existing flow policies learn instantaneous velocities and require multi-step iterative sampling (lipman et al., 2023; park et al., 2025; bharadhwaj et al., 2024), mfp instead learns the mean velocity field (geng et al., 2025a). this design enables a direct, single-step mapping from a base gaus... - **Candidate:** To enable high-quality one-step generation, we reinterpret the velocity field  $v(a, t; s)$  in flow matching as the instantaneous velocity, and instead propose to model the average velocity, which directly connects any two arbitrary time steps. specifically, we define the average velocity over an interval...

Evidence 2 - **Rationale:** The mathematical formulation for mean/average velocity is identical in both papers, demonstrating that this specific technical approach was already established. - **Original:** Unlike standard flow policies that learn the instantaneous velocity field  $v(a(t), t, s)$ , we center on the mean velocity field  $u(a(t), t, r, s)$ , modeling the mean velocity over any given time interval  $[t, r]$  as  $u(a(t), t, r, s) \triangleq \int_t^r v(a(\tau), \tau, s) d\tau$ . - **Candidate:** specifically, we define the average velocity over an interval  $[r, t]$  as  $u(a, r, t; s) \triangleq \int_r^t v(a(\tau), \tau, s) d\tau$ , (10) representing the total displacement across the interval divided by its duration.

Evidence 3 - **Rationale:** Both papers claim to enable one-step action generation by modeling mean/average velocity fields, with Diffusion Q Learning[18] explicitly stating this approach and its benefits prior to the original paper's submission. - **Original:** we propose a new flow-based policy, namely mean flow policy (mfp), that enables fastest one-step action generation. by modeling the mean velocity field, mfp retains the expressiveness of generative policies while eliminating multi-step sampling overhead. - **Candidate:** we introduce one-step flow q-learning (ofql), a novel framework that enables effective one-step action generation during both training and inference, without auxiliary modules or distillation. ofql reformulates the dql policy within the flow matching (fm) paradigm but departs from conventional fm by...

---

### 5. Time-Unified Diffusion Policy with Action Discrimination for Robotic Manipulation

URL: [View paper](#)

#### Brief Assessment

Time Unified Diffusion[45] focuses on diffusion-based policies for robotic manipulation with time-unified denoising, not flow-based policies modeling mean velocity fields as in the original paper.

---

### 6. Improved Mean Flows: On the Challenges of Fastforward Generative Models

URL: [View paper](#)

#### Brief Assessment

Improved Mean Flows[47] focuses on image generation using mean velocity fields, not reinforcement learning or action generation for robotic control tasks.

---

### 7. Towards High-Order Mean Flow Generative Models: Feasibility, Expressivity, and Provably Efficient Criteria

URL: [View paper](#)

#### Brief Assessment

High Order Mean[46] focuses on theoretical foundations of second-order mean flow for generative modeling (incorporating average acceleration fields), not on RL policy design for action generation. The candidate addresses circuit complexity and attention approximation in generative models, while the original paper proposes a practical RL policy with instantaneous velocity constraints.

---

## 8. One-Step Generative Channel Estimation via Average Velocity Field

URL: [View paper](#)

### Brief Assessment

Generative Channel Estimation[48] applies mean velocity field modeling to wireless channel estimation, not reinforcement learning action generation. The domains and applications are fundamentally different.

---

## 9. OM2P: Offline multi-agent mean-flow policy

URL: [View paper](#)

### Prior Art Analysis

OM2P[44] demonstrates that mean flow policies for one-step action generation were already applied in offline multi-agent reinforcement learning prior to the original paper's submission. Both papers model mean velocity fields instead of instantaneous velocities to enable direct single-step mapping from Gaussian noise to action distributions. The candidate paper explicitly states it introduces mean-flow model as policy networks for offline MARL to achieve one-step action generation, and provides the same mathematical formulation for mean velocity as  $u(a(t), t, r, s) = (1/(r-t)) \int [t \text{ to } r] v(a(\tau), \tau, s) d\tau$ . This directly challenges the novelty claim that the original authors were first to introduce this approach.

### Evidence

Evidence 1 - **Rationale:** Both papers claim to introduce mean-flow policies for one-step action generation. The candidate explicitly states this was done for offline MARL, demonstrating prior work exists in this direction. - **Original:** we propose mean flow policy (mfp), a new generative policy function that models the mean velocity field to achieve the fastest one-step action generation - **Candidate:** we propose om 2p ( offline multi-agent mean-flow policy )-a novel framework that seamlessly incorporates mean-flow model as policy networks for offline marl. om 2p redefines the training paradigm of mean-flow-based policy networks via generalized timestep sampling and derivative-free target mean-vel...

Evidence 2 - **Rationale:** Both papers describe the same core innovation: replacing instantaneous velocity fields with mean velocity fields to enable one-step action generation, eliminating multi-step sampling. - **Original:** while existing flow policies learn instantaneous velocities and require multi-step iterative sampling (lipman et al., 2023; park et al., 2025; bharadhwaj et al., 2024), mfp instead learns the mean velocity field (geng et al., 2025a). this design enables a direct, single-step mapping from a base gaus... - **Candidate:** recent advances in flow-based models, particularly mean-flow model [12], have demonstrated the potential for efficient one-step sampling by replacing instantaneous velocity fields with their averaged counterparts, enabling fast and high-quality generation in image synthesis. inspired by mean-flow mo...

Evidence 3 - **Rationale:** Both papers use the same mean-flow identity formulation for training the mean velocity field, showing identical mathematical approaches to the problem. - **Original:** the mean velocity model  $u$ , we first multiply both sides of eq. (6) by  $(r - t)$  and then differentiate with respect to  $t$  ( treating  $r$  as independent of  $t$  ), which yields the mean flow identity:  $-u(a(t), t, r, s) + (r - t) \frac{d}{dt} u(a(t), t, r, s) = -v(a(t), t, s)$  - **Candidate:** to model the mean velocity, we adopt a neural network  $u_{\theta_{pxr}, r, tq}$  trained to satisfy the mean-flow identity (equals to equation (9) in [12]) via:  $\| \theta_{q} \|_2 \leq \epsilon$   $\| u_{\theta_{pxr}, r, tq} - v_{\theta_{pxr}, r, tq} \|_2 \leq \epsilon$  where  $x_r = [p_1 \dots p_n] \in \mathbb{R}^{n \times 1}$ , with  $x_0 = [np_0, idq, x_1, \dots, p_{pxq}]$ , and  $r, t, u, v, \theta_{pxr}, 1sq$ . the target velocity i...

---

## 10. Flowpolicy: Enabling fast and robust 3d flow-based policy via consistency flow matching for robot manipulation

URL: [View paper](#)

### Brief Assessment

Flowpolicy Consistency[3] focuses on consistency flow matching for robotic manipulation with 3D visual representations, not on modeling mean velocity fields as described in the original paper's contribution.

---

## Contribution 2: Instantaneous Velocity Constraint as boundary condition

**Description:** The authors propose IVC, a training objective that pairs average velocity loss with instantaneous velocity loss at interval start points. They theoretically prove this constraint acts as a necessary boundary condition for the mean flow ODE, eliminating solution multiplicity and improving learning accuracy with negligible computational overhead.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

## 1. SpecGuard: Specification aware recovery for robotic autonomous vehicles from physical attacks

URL: [View paper](#)

### Brief Assessment

SpecGuard[39] focuses on recovery control for robotic autonomous vehicles under physical attacks using deep reinforcement learning with mission specification compliance. It does not address boundary conditions or velocity constraints for training flow-based policies in reinforcement learning, which is the core contribution of the original paper.

---

## 2. Control-barrier-function-based design of gradient flows for constrained nonlinear programming

URL: [View paper](#)

### Brief Assessment

Control Barrier Gradient[38] addresses constrained nonlinear optimization using control barrier functions and gradient flows, not reinforcement learning policy training with flow-based models and velocity constraints.

---

## 3. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations

URL: [View paper](#)

### Brief Assessment

Hidden Fluid Mechanics[35] focuses on extracting velocity and pressure fields from flow visualizations using Navier-Stokes equations for fluid dynamics, not on training flow-based policies for reinforcement learning with boundary conditions.

---

## 4. Uncertainty-aware constraint inference in inverse constrained reinforcement learning

URL: [View paper](#)

### Brief Assessment

Uncertainty Constraint Inference[36] focuses on constraint inference in inverse constrained RL using distributional methods and flow-based data augmentation, not on boundary conditions for flow-based policy training in standard RL.

---

## 5. Achieving Safe Control Online through Integration of Harmonic Control Lyapunov-Barrier Functions with Unsafe Object-Centric Action Policies

URL: [View paper](#)

### Brief Assessment

Harmonic Control Lyapunov[42] addresses safe robot control through harmonic control lyapunov-barrier functions for spatial constraints, not velocity field training for flow-based policies in RL.

---

## 6. A numerical study of fish adaption behaviors in complex environments with a deep reinforcement learning and immersed boundaryâ€”lattice Boltzmann method

URL: [View paper](#)

### Brief Assessment

Fish Adaption Reinforcement[37] focuses on fish swimming behaviors using deep reinforcement learning coupled with fluid dynamics simulation (IB-LBM), not on training objectives or boundary conditions for flow-based policies in reinforcement learning.

---

## 7. Flow-Based Policy for Online Reinforcement Learning

URL: [View paper](#)

### Brief Assessment

Flow Online Reinforcement[40] focuses on Wasserstein-2-regularized optimization for online RL with flow-based policies, not on boundary conditions for mean velocity ODEs. The candidate addresses policy-value alignment through distributional constraints, while the original addresses ODE solution multiplicity through instantaneous velocity constraints at interval boundaries.

---

## 8. SAC Flow: Sample-Efficient Reinforcement Learning of Flow-Based Policies via Velocity-Reparameterized Sequential Modeling

URL: [View paper](#)

### Brief Assessment

SAC Flow[34] focuses on stabilizing gradient flow through sequential model reparameterizations (GRU-gated and Transformer-decoded velocities) rather than boundary conditions for mean velocity ODEs. The candidate addresses gradient pathologies in multi-step rollouts, not the solution multiplicity problem that IVC targets.

---

## 9. A dynamic programming approach for optimizing train speed profiles with speed restrictions and passage points

URL: [View paper](#)

### Brief Assessment

Train Speed Optimization[41] addresses train speed profile optimization with speed restrictions and passage points, which is a completely different domain from reinforcement learning flow-based policies. The candidate focuses on railway engineering optimization problems, not on training objectives or boundary conditions for neural network-based policy learning in RL.

---

## 10. Flowpolicy: Enabling fast and robust 3d flow-based policy via consistency flow matching for robot manipulation

URL: [View paper](#)

### Brief Assessment

Flowpolicy Consistency[3] employs consistency constraints on velocity fields but does not propose IVC as a boundary condition pairing average and instantaneous velocity losses at interval start points.

---

## Contribution 3: State-of-the-art performance with substantial speedup

**Description:** The authors demonstrate that their method achieves top success rates on Robomimic and OGBench benchmarks while delivering significant improvements in training and inference speed compared to multi-step flow policy baselines, highlighting practical applicability for real-time robotic control.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

### 1. MP1: MeanFlow Tames Policy Learning in 1-step for Robotic Manipulation

URL: [View paper](#)

#### Prior Art Analysis

MP1 MeanFlow[4] demonstrates that achieving state-of-the-art performance with substantial speedup using flow policies on robotic manipulation benchmarks was already accomplished prior to the original paper. The candidate reports achieving 78.9% average success rate with 6.8ms inference time, outperforming diffusion baselines by significant margins. The candidate explicitly compares against the same baseline methods (DP3, FlowPolicy) and demonstrates both superior success rates and faster inference speeds on manipulation benchmarks, establishing this combination of performance and efficiency before the original work.

#### Evidence

Evidence 1 - **Rationale:** The candidate explicitly demonstrates the same achievement: state-of-the-art success rates combined with substantial speed improvements over flow-based baselines. - **Original:** our mfp achieves state-of-the-art success rates across several challenging robotic manipulation tasks from robomimic and ogbench. it also delivers substantial improvements in training and inference speed over existing flow-based policy baselines - **Candidate:** mp1 achieves superior average task success rates, outperforming dp3 by 10.2% and flowpolicy by 7.3%. its average inference time is only 6.8 ms-19x faster than dp3 and nearly 2x faster than flowpolicy

Evidence 2 - **Rationale:** Both papers claim state-of-the-art performance on manipulation benchmarks with substantial improvements over the same baseline methods. - **Original:** as shown in figure 1 and table 3, mfp achieves state-of-the-art success rates while delivering substantial speed-ups in training and per-step inference over strong flow-policy baselines, on average across both suites - **Candidate:** tab. 1 demonstrates that the mp1 achieves sota performance across all sub-tasks. the overall average success rate reaches 78.9%±2.1%, which significantly outperforms the previous best method, flowpolicy, at 71.6% ±3.5%. compared to existing approaches, our method yields a 10.2% improvement over dp3 ...

Evidence 3 - **Rationale:** Both papers demonstrate the combination of highest success rates with fastest speeds, establishing this performance-efficiency trade-off before the original work. - **Original:** our mfp achieved the highest success rate and fastest training speed on a single a100 gpu - **Candidate:** mp1 achieves sota inference speed across all sub-tasks, with an average latency of just 6.8 ms-nearly 2x faster than the best flowpolicy (which relies on consistency constraints for its 1-nfe sampling ) and nearly 14x faster than diffusion-based methods

Evidence 4 - **Rationale:** The candidate's contribution statement directly parallels the original's claim, demonstrating the same achievement on robotic manipulation benchmarks with both performance and speed improvements. - **Original:** mfp achieves state-of-the-art success rates on two challenging robotic manipulation benchmarks: robomimic and ogbench. moreover, our approach provides a substantial speedup in both training and inference over existing flow-policy baselines - **Candidate:** extensive experiments on adroit and meta-world simulation benchmarks, as well as real-world tasks, show that mp1 surpasses diffusion and flow-based baselines in both success rate and speed

---

## 2. Vfp: Variational flow-matching policy for multi-modal robot manipulation

URL: [View paper](#)

### Brief Assessment

Variational Flow Matching[30] focuses on multi-modal robot manipulation using variational inference and optimal transport, not on general RL frameworks or the specific flow policy training methods proposed in the original paper. The candidate addresses different technical challenges (multi-modality in imitation learning) rather than the original's focus on mean velocity fields and instantaneous velocity constraints for online RL.

---

## 3. FlowRAM: Grounding Flow Matching Policy with Region-Aware Mamba Framework for Robotic Manipulation

URL: [View paper](#)

### Brief Assessment

FlowRAM[7] focuses on robotic manipulation with region-aware perception and flow matching for keyframe pose generation, while the original paper addresses general RL policy learning with mean flow for one-step action generation across different benchmarks (Robomimic and OGBench). The technical approaches and application domains differ substantially.

---

## 4. DM1: MeanFlow with Dispersive Regularization for 1-Step Robotic Manipulation

URL: [View paper](#)

### Brief Assessment

DM1 Dispersive[12] focuses on preventing representation collapse in flow-based policies through dispersive regularization, a different technical approach than the original paper's instantaneous velocity constraint (IVC) for mean flow policies. The candidate addresses a distinct problem (representation collapse) rather than challenging the novelty of achieving fast inference through mean velocity field modeling.

---

## 5. Can we detect failures without failure data? uncertainty-aware runtime failure detection for imitation learning policies

URL: [View paper](#)

### Brief Assessment

Runtime Failure Detection[33] focuses on detecting policy failures during execution without failure data, not on achieving fast inference for flow policies on robotic benchmarks. The candidate addresses a different problem (failure detection) than the original's contribution (policy performance and speed).

---

## 6. Genie Envisioner: A Unified World Foundation Platform for Robotic Manipulation

URL: [View paper](#)

### Brief Assessment

Genie Envisioner[31] focuses on video-generative world models for robotic manipulation using diffusion models and flow-matching decoders, not on flow-based policy learning for direct action generation in RL benchmarks.

---

## 7. 3D FlowMatch Actor: Unified 3D Policy for Single- and Dual-Arm Manipulation

URL: [View paper](#)

### Brief Assessment

FlowMatch Actor[8] focuses on 3D visual policy learning with flow matching for robotic manipulation, while the original paper addresses general RL policy learning with mean flow for state-based control. The candidate's speedup comes from architectural optimizations in 3D vision-based policies, not from one-step flow generation methods.

---

## 8. General Flow as Foundation Affordance for Scalable Robot Learning

URL: [View paper](#)

### Brief Assessment

General Flow Affordance[32] focuses on zero-shot human-to-robot skill transfer using 3D flow prediction from human videos, not on achieving fast inference speeds in robotic manipulation benchmarks through flow-based RL policies.

---

## 9. : A Vision-Language-Action Flow Model for General Robot Control

URL: [View paper](#)

### Brief Assessment

Vision Language Action[29] focuses on vision-language-action models for general robot control across diverse embodiments, not on flow policy inference speed optimization for robotic manipulation benchmarks like the original paper's MFP method.

---

## 10. Flowpolicy: Enabling fast and robust 3d flow-based policy via consistency flow matching for robot manipulation

URL: [View paper](#)

### Prior Art Analysis

Flowpolicy Consistency[3] demonstrates that flow-based policies can achieve fast inference speeds on robotic manipulation benchmarks before the original paper's publication. The candidate reports achieving real-time inference (19.9ms average) with competitive success rates across 37 tasks from Adroit and Metaworld benchmarks, representing a 7x speedup over baseline methods. This establishes prior work showing flow policies achieving both high performance and substantial speed improvements on robotic manipulation tasks.

### Evidence

Evidence 1 - **Rationale:** Both papers claim achieving state-of-the-art success rates with improved inference speed on robotic manipulation benchmarks, establishing that this combination was demonstrated in prior work. - **Original:** mfp achieves state-of-the-art success rates across several challenging robotic manipulation tasks from robomimic and ogbench. it also delivers substantial improvements in training and inference speed over existing flow-based policy baselines - **Candidate:** flowpolicy's overall success rate is 1.3 % higher than dp3 and 2.6 % higher than simple dp3 with the same expert demonstration. the results show that flowpolicy can achieve competitive results with state-of-the-art methods in just one step of inference

Evidence 2 - **Rationale:** The candidate explicitly demonstrates achieving competitive success rates with substantial speedup on robotic manipulation benchmarks, which directly overlaps with the original paper's claimed contribution. - **Original:** our mfp achieves state-of-the-art success rates on two challenging robotic manipulation benchmarks: robomimic and ogbench. moreover, our approach provides a substantial speedup in both training and inference over existing flow-policy baselines - **Candidate:** we validate the effectiveness of flowpolicy in adroit and metaworld, demonstrating a 7x increase in inference speed while maintaining competitive average success rates compared to state-of-the-art methods

---

## Appendix: Text Similarity Detection

Textual similarity detection checked 27 papers and found 5 similarity segment(s) across 2 paper(s).

The following **2 paper(s)** were detected to have high textual similarity with the original paper. These may represent different versions of the same work, duplicate submissions, or papers with substantial textual overlap. Readers are advised to verify these relationships independently.

### 1. MP1: MeanFlow Tames Policy Learning in 1-step for Robotic Manipulation

**Detected in:** Core Task (sibling), Contribution: contribution\_1, Contribution: contribution\_3

△ **Note:** This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

---

### 2. OM2P: Offline multi-agent mean-flow policy

**Detected in:** Contribution: contribution\_1

△ **Note:** This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

---

## References

- [0] Mean Flow Policy with Instantaneous Velocity Constraint for One-step Action Generation [View paper](#)
- [1] Adaflow: Imitation learning with variance-adaptive flow-based policies [View paper](#)
- [2] Flow as the cross-domain manipulation interface [View paper](#)
- [3] Flowpolicy: Enabling fast and robust 3d flow-based policy via consistency flow matching for robot manipulation [View paper](#)
- [4] MP1: MeanFlow Tames Policy Learning in 1-step for Robotic Manipulation [View paper](#)
- [5] Flow-Based Single-Step Completion for Efficient and Expressive Policy Learning [View paper](#)
- [6] Manifold: A general robot manipulation policy via consistency flow training [View paper](#)
- [7] FlowRAM: Grounding Flow Matching Policy with Region-Aware Mamba Framework for Robotic Manipulation [View paper](#)
- [8] 3D FlowMatch Actor: Unified 3D Policy for Single- and Dual-Arm Manipulation [View paper](#)
- [9] FlowDreamer: A RGB-D World Model with Flow-based Motion Representations for Robot Manipulation [View paper](#)
- [10] FreqPolicy: Efficient Flow-based Visuomotor Policy via Frequency Consistency [View paper](#)
- [11] FBI: Learning Dexterous In-hand Manipulation with Dynamic Visuotactile Shortcut Policy [View paper](#)
- [12] DM1: MeanFlow with Dispersive Regularization for 1-Step Robotic Manipulation [View paper](#)
- [13] ReinFlow: Fine-tuning flow matching policy with online reinforcement learning [View paper](#)
- [14] Towards a Generalizable Bimanual Foundation Policy via Flow-based Video Prediction [View paper](#)
- [15] FabricFlowNet: Bimanual Cloth Manipulation with a Flow-based Policy [View paper](#)
- [16] FRMD: Fast Robot Motion Diffusion with Consistency-Distilled Movement Primitives for Smooth Action Generation [View paper](#)
- [17] Dense-Jump Flow Matching with Non-Uniform Time Scheduling for Robotic Policies: Mitigating Multi-Step Inference Degradation [View paper](#)
- [18] Revisiting Diffusion Q-Learning: From Iterative Denoising to One-Step Action Generation [View paper](#)
- [19] Streaming Flow Policy: Simplifying diffusion/flow-matching policies by treating action trajectories as flow trajectories [View paper](#)
- [20] One-step flow policy mirror descent [View paper](#)
- [21] One-Step Generative Policies with Q-Learning: A Reformulation of MeanFlow [View paper](#)
- [22] Decision Flow Policy Optimization [View paper](#)
- [23] Compose Your Policies! Improving Diffusion-based or Flow-based Robot Policies via Test-time Distribution-level Composition [View paper](#)
- [24] SeFA-Policy: Fast and Accurate Visuomotor Policy Learning with Selective Flow Alignment [View paper](#)
- [25] L1 Sample Flow for Efficient Visuomotor Learning [View paper](#)
- [26] Fast Flow-based Visuomotor Policies via Conditional Optimal Transport Couplings [View paper](#)
- [27] Fast and Accurate Visuomotor Imitation Learning via 2D Consistency Flow Matching Policy [View paper](#)
- [28] RecFlow Policy: Fast and Accurate Visuomotor Policy Learning via Rectified Action Flow [View paper](#)
- [29] : A Vision-Language-Action Flow Model for General Robot Control [View paper](#)
- [30] Vfp: Variational flow-matching policy for multi-modal robot manipulation [View paper](#)
- [31] Genie Envisioner: A Unified World Foundation Platform for Robotic Manipulation [View paper](#)
- [32] General Flow as Foundation Affordance for Scalable Robot Learning [View paper](#)
- [33] Can we detect failures without failure data? uncertainty-aware runtime failure detection for imitation learning policies [View paper](#)
- [34] SAC Flow: Sample-Efficient Reinforcement Learning of Flow-Based Policies via Velocity-Reparameterized Sequential Modeling [View paper](#)
- [35] Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations [View paper](#)
- [36] Uncertainty-aware constraint inference in inverse constrained reinforcement learning [View paper](#)
- [37] A numerical study of fish adaption behaviors in complex environments with a deep reinforcement learning and immersed boundary lattice Boltzmann method [View paper](#)
- [38] Control-barrier-function-based design of gradient flows for constrained nonlinear programming [View paper](#)
- [39] SpecGuard: Specification aware recovery for robotic autonomous vehicles from physical attacks [View paper](#)
- [40] Flow-Based Policy for Online Reinforcement Learning [View paper](#)
- [41] A dynamic programming approach for optimizing train speed profiles with speed restrictions and passage points [View paper](#)
- [42] Achieving Safe Control Online through Integration of Harmonic Control Lyapunov-Barrier Functions with Unsafe Object-Centric Action Policies [View paper](#)
- [43] Mean Flows for One-step Generative Modeling [View paper](#)
- [44] OM2P: Offline multi-agent mean-flow policy [View paper](#)

- [45] Time-Unified Diffusion Policy with Action Discrimination for Robotic Manipulation [View paper](#)
- [46] Towards High-Order Mean Flow Generative Models: Feasibility, Expressivity, and Provably Efficient Criteria [View paper](#)
- [47] Improved Mean Flows: On the Challenges of Fastforward Generative Models [View paper](#)
- [48] One-Step Generative Channel Estimation via Average Velocity Field [View paper](#)