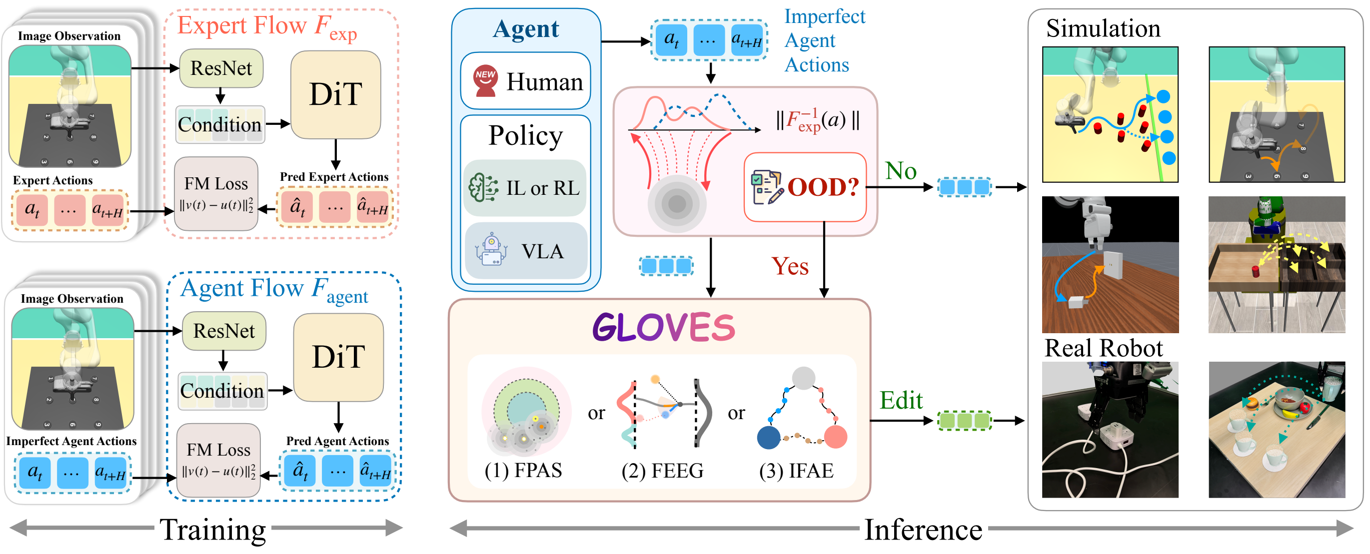
GLOVES: Flow-based Policy Adaptation without Policy Updates

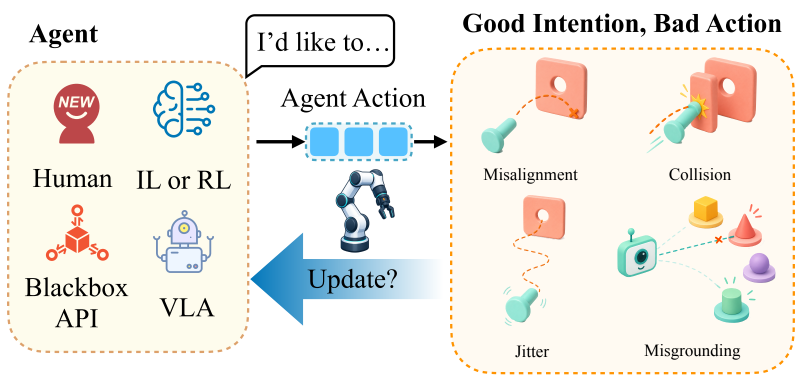
Luzhe Sun[*1], *Jingtian Ji*[*1], *Haoran Chen*[1], *Jiawei Zhou*[2], *Matthew R. Walter*[1]

[1]*Toyota Technological Institute at Chicago*, [2]*Stony Brook University*



Why policy adaptation?

Agents often contain useful task intent but produce imperfect executions; policy adaptation corrects their proposed actions online without requiring policy updates, reward design, or access to dynamics.

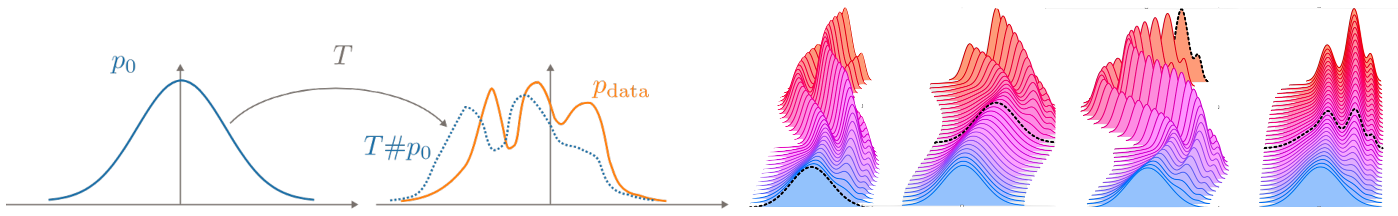


What is flow matching?

Flow matching define a velocity field by ODE[1]:

$$\frac{d\phi_t(x)}{dt} = v_t\left(\phi_t(x)\right)$$

$\phi_t(x)$: flow map, position of particle at time t $\qquad$ v_t : velocity field, how particle moves $t = 0 \rightarrow 1$



Out of distribution detection

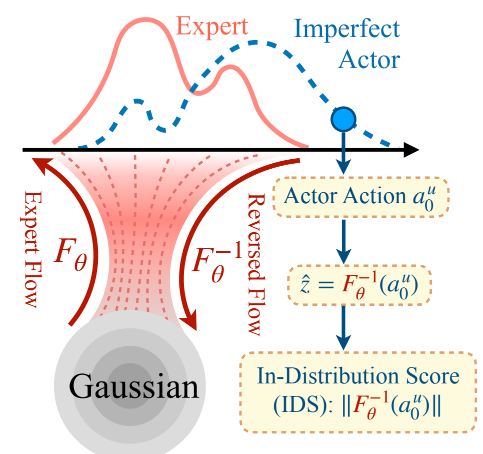
Given a trained flow $F_\theta(\cdot)$ and
a non-conformity score function score $= \hat{z} = F_\theta^{-1}(x)$ [2] and
a calibration set $D_{\text{cal}} = \{(x_i)\}_{i=1}^{n}$

$$p\text{-value} = \frac{|\, j \in \{1,\ldots,n\} : F_\theta^{-1}(x_j) \geq F_\theta^{-1}(x_{\text{test}})\,| + 1}{n+1}$$

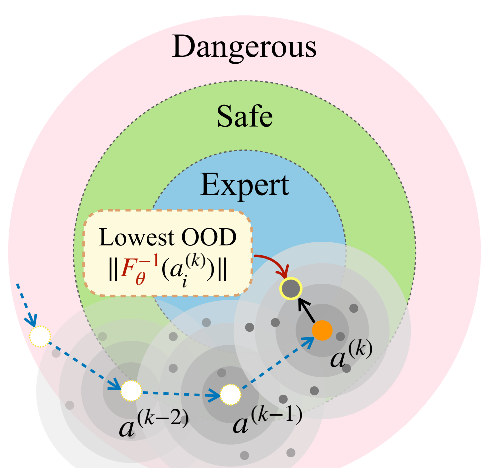
Larger nonconformity score $\rightarrow$ smaller p-value $\rightarrow$ more OOD.

In distribution $C_\alpha(x_{\text{test}}) = \{p\text{-value}(x_{\text{test}}) > \alpha\}$

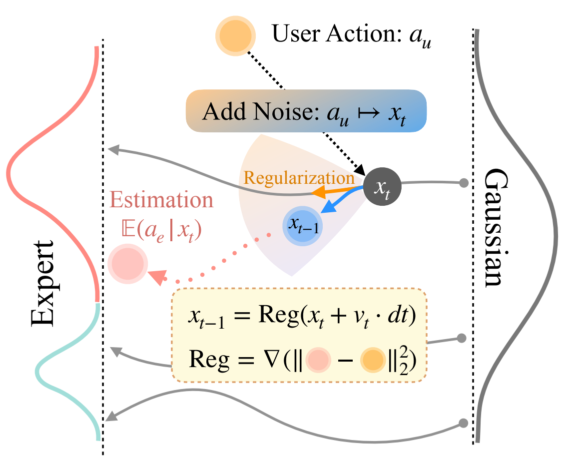
All we need for conformal prediction is expert flow.



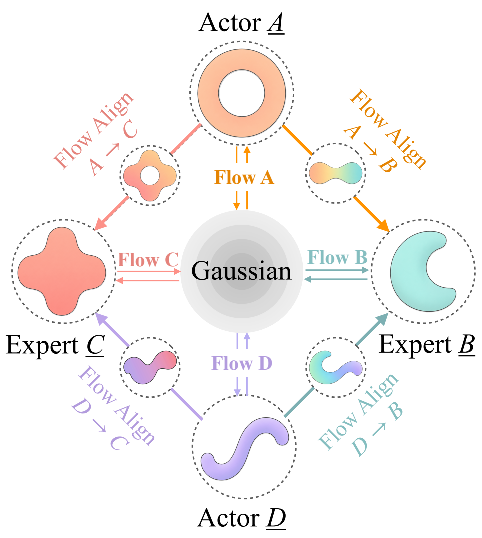
GLOVES: Policy Adaptation without Policy Updates



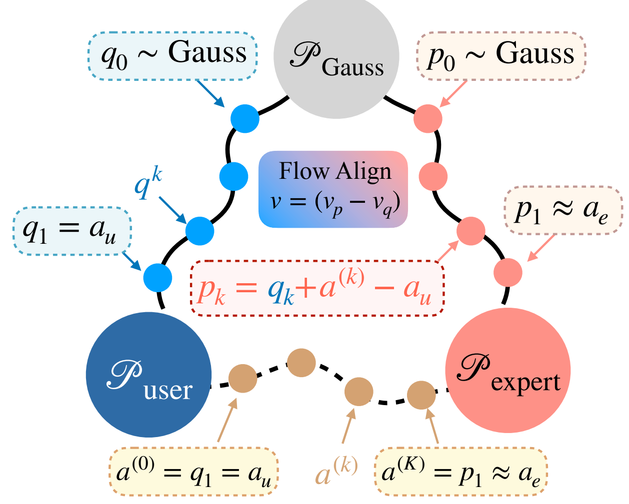
(a) FPAS
Flow-Prior Action Sampling



(b) FEEG
Flow Editing with Energy Guidance

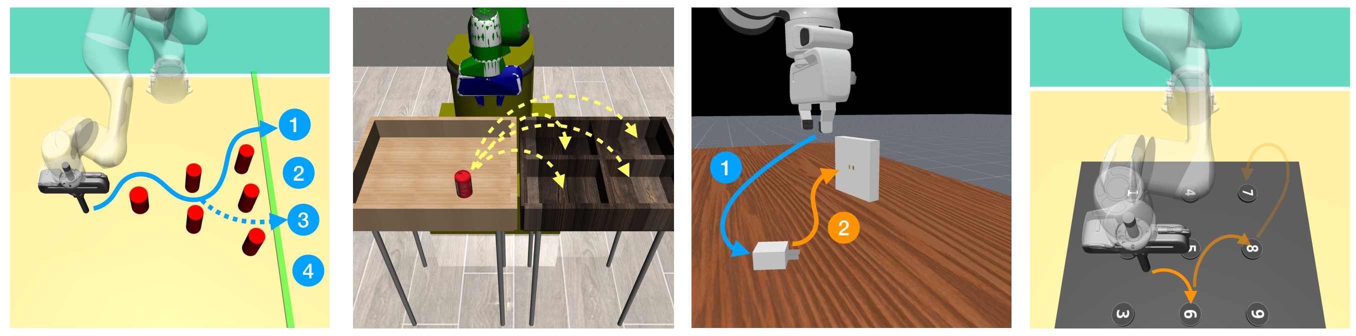


Preliminary: Flow Align[3]



(c) IFAE
Inversion-Free Action-chunk Editing

Experiments



(a) Slalom $\qquad$ (b) Can Transfer $\qquad$ (c) Charger Insertion $\qquad$ (d) Key Pad

(e) Real Robot Charger Insertion $\qquad$ (f) Cup Serving with Fluid

Agent Design

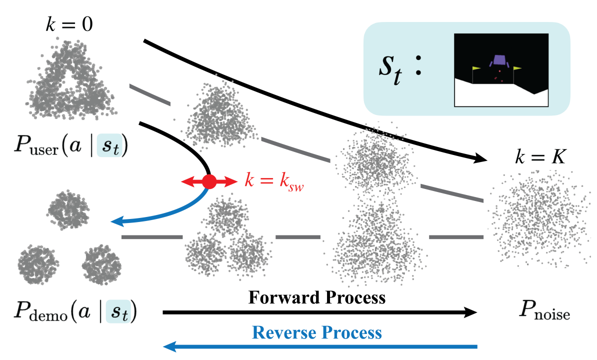
IL Agent: A goal-conditioned imitation policy is trained from expert demonstrations and then perturbed with noise, delay, slowing, or bias. This creates controlled imperfect agents with known failure modes.

$$\tilde{\mathbf{a}}_k^{\text{laggy}} = \begin{cases} \tilde{\mathbf{a}}_{k-1}, & \text{with probability } p_{\text{lag}}, \\ \mathbf{a}_k, & \text{otherwise}. \end{cases} \qquad \tilde{\mathbf{a}}_k^{\text{shift}} = \begin{cases} \mathbf{a}_k + \mathbf{1}_T(\delta \mathbf{e}_j)^\top, & \text{with probability } p_{\text{shift}}, \\ \mathbf{a}_k, & \text{otherwise}. \end{cases}$$

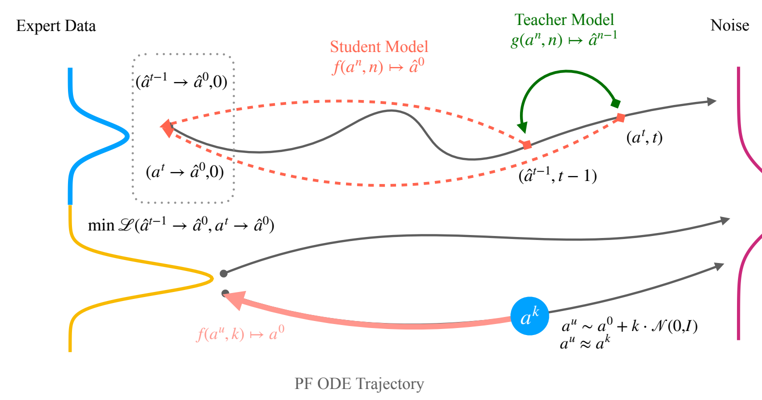
$$\tilde{\mathbf{a}}_k^{\text{slow}} = \begin{cases} \alpha\,\mathbf{a}_k, & \text{with probability } p_{\text{slow}}, \quad 0 < \alpha < 1 \\ \mathbf{a}_k, & \text{otherwise}, \end{cases} \qquad \tilde{\mathbf{a}}_k^{\text{noised}} = \mathbf{a}_k + \boldsymbol{\epsilon}_k, \qquad \boldsymbol{\epsilon}_k \sim \mathcal{N}(0, \sigma^2 I).$$

VLA Agent: A finetuned vision-language-action model(FLOWER[4]) proposes action chunks from images and language instructions. GLOVES improves its execution online while keeping the VLA frozen.

Baselines



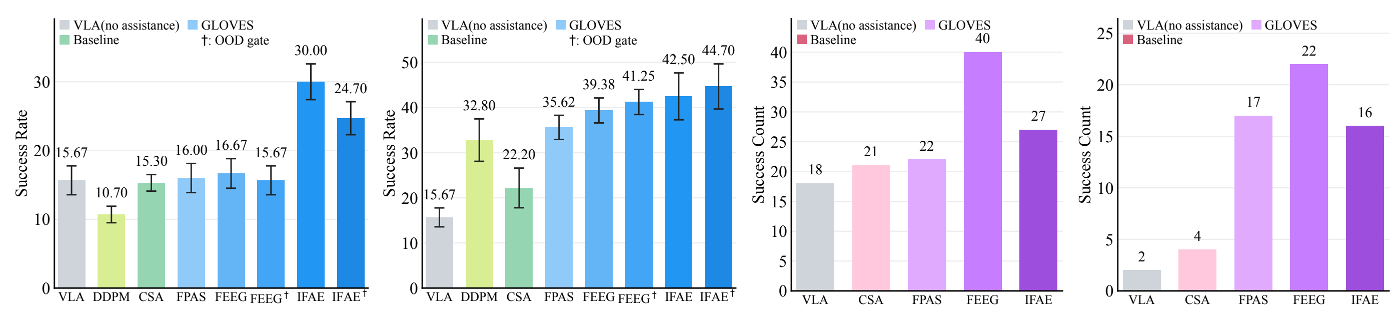
DDPM for shared autonomy[5]



Consistency model for shared autonomy(CSA)[6]

Results

Task	Agent	Agent Performance	DDPM	CSA	FPAS	FEEG w/o OOD	FEEG w/ OOD	IFAE w/o OOD	IFAE w/ OOD
Slalom	laggy	38.7±5.8	15.3±2.2	**43.6±5.0**	**41.0±7.6**	31.7±3.8	31.0±3.7	39.4±5.4	39.3±5.3
	noised	40.0±7.1	33.3±3.3	42.7±5.3	46.0±6.1	46.7±5.8	45.7±5.8	**53.7±5.3**	**53.6±5.4**
	slow	55.7±5.1	35.7±3.3	42.3±4.0	60.0±5.0	49.3±5.6	49.3±5.6	**74.6±4.6**	**74.0±4.4**
	shift	28.0±2.8	15.0±2.3	**69.7±4.0**	47.7±7.3	40.0±3.5	36.7±4.0	65.0±4.7	**70.0±4.7**
Can	laggy	48.1±8.2	53.8±5.3	51.5±5.3	53.8±7.1	55.0±6.3	53.1±6.5	**66.6±5.1**	**81.9±4.4**
	noised	50.9±5.6	1.6±1.4	**60.3±5.0**	48.4±3.2	52.5±5.7	50.6±5.8	**55.9±5.2**	55.6±5.0
	slow	23.1±6.0	62.2±5.0	70.9±4.8	29.4±4.6	62.8±3.5	60.9±4.8	**72.2±4.7**	**71.6±4.7**
	shift	54.7±3.9	80.3±4.2	52.5±5.5	57.2±3.2	**93.4±3.3**	**92.2±3.2**	56.9±5.4	58.1±5.6
Keypad	laggy	54.0±8.1	57.3±2.3	64.0±2.3	62.3±5.3	61.3±6.3	62.3±6.6	**82.0±1.7**	**81.7±1.7**
	noised	34.7±5.6	47.7±2.3	25.7±1.7	36.0±6.1	**51.3±7.8**	**50.7±6.9**	46.0±2.7	44.8±2.4
	slow	52.7±6.2	81.3±2.0	92.0±1.0	94.0±2.4	**95.0±2.6**	**96.0±2.1**	79.3±2.0	88.0±2.0
	shift	52.7±5.6	66.0±2.0	68.0±2.0	68.3±5.1	**82.0±3.5**	**80.7±4.0**	61.4±1.7	57.3±2.0
Charger	laggy	38.5±7.0	0.0±0.0	40.0±7.0	41.0±6.8	39.0±7.8	**41.0±5.6**	40.0±7.0	**44.5±7.0**
	noised	33.0±5.7	0.0±0.0	48.0±7.0	61.5±7.7	**68.5±4.4**	**73.0±5.3**	65.5±6.5	66.4±6.9
	slow	45.0±5.3	0.0±0.0	**49.9±7.0**	44.5±3.7	46.0±5.5	**47.5±5.1**	46.0±7.0	44.5±7.0
	shift	36.0±5.0	0.0±0.0	60.5±6.5	49.0±4.8	**68.0±5.7**	**64.5±4.3**	36.0±7.0	40.4±6.6



(a) Slalom $\qquad$ (b) CAN $\qquad$ (c) Real Charger $\qquad$ (d) Cup Serve

[1] Y. Lipman, R. T. Q. Chen, H. Ben-Hamu, M. Nickel, and M. Le. Flow matching for generative modeling. ICLR 2023
[2] C. Xu, T. K. Nguyen, E. Dixon, C. Rodriguez, P. Miller, R. Lee, P. Shah, R. Ambrus, H. Nishimura, and M. Itkina. Can we detect failures without failure data? Uncertainty-aware runtime failure detection for imitation learning policies. RSS 2025
[3] J. Kim, Y. Hong, J. Park, and J. C. Ye. FlowAlign: Trajectory-regularized, inversion-free flow-based image editing. ICLR 2026
[4] M. Reuss, H. Zhou, M. Ruehle, O. E. Yagmurlu, F. Otto, and R. Lioutikov. FLOWER: Democratizing generalist robot policies with efficient vision-language-action flow policies.
[5] T. Yoneda, L. Sun, G. Yang, B. Stadie, and M. Walter. To the noise and back: Diffusion for shared autonomy. RSS 2023
[6] L. Sun, J. Ji, X. Tan, and M. Walter. FlashBack: Consistency model-accelerated shared autonomy. CoRL 2025