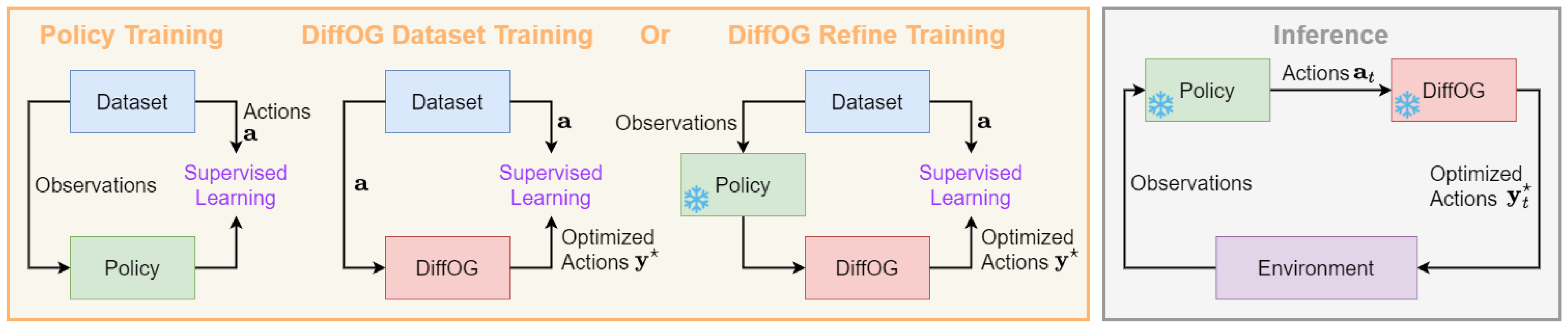
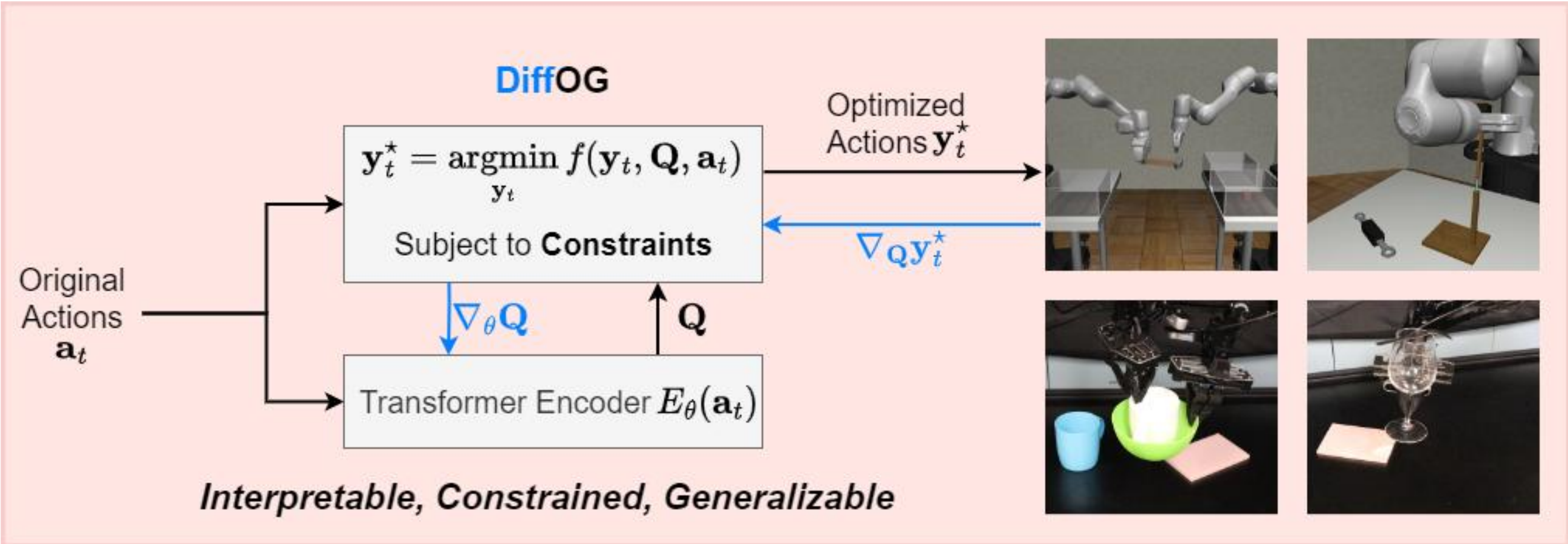
DiffOG: Differentiable Policy Trajectory Optimization with Generalizability

Zhengtong Xu, Zichen Miao, Qiang Qiu, Zhe Zhang, Yu She
{xu1703, miaoz, qqiu, zhan5111, shey}@purdue.edu
PURDUE UNIVERSITY

Method

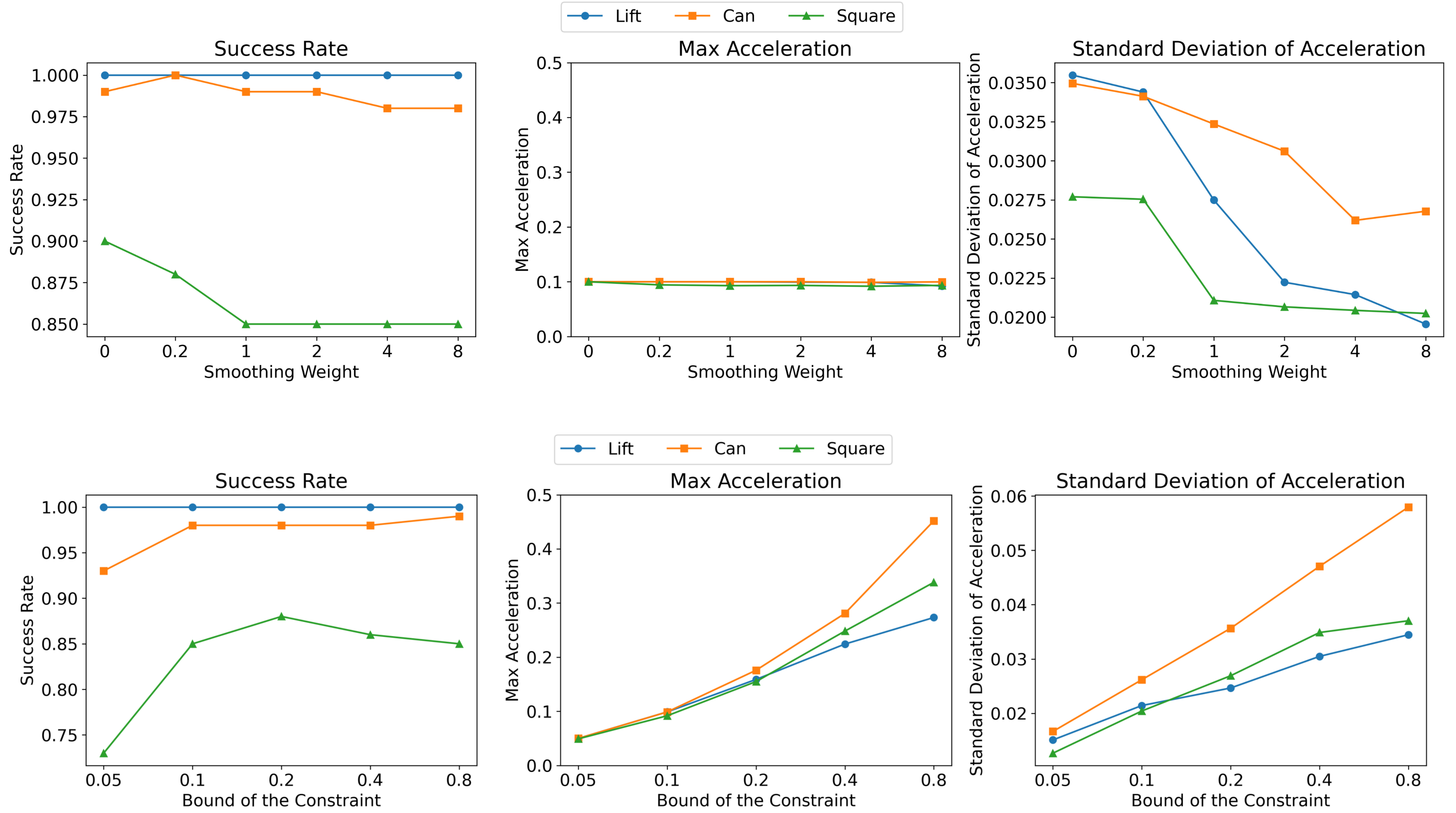
We introduce DiffOG, a learning-based trajectory optimizer that enhances visuomotor policies by integrating a differentiable transformer-based optimization layer. DiffOG produces smoother, constraint-compliant, and more interpretable action trajectories.





Adjustability of Trajectory

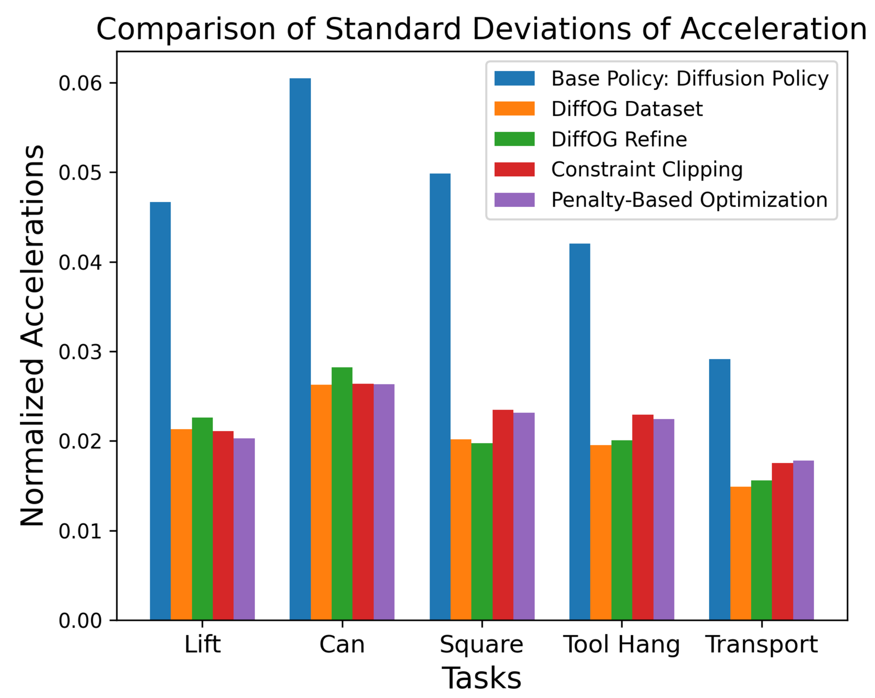
Visuomotor policies typically lack mechanisms for directly influencing the properties of the generated actions through parameter adjustment. DiffOG offers interpretability, enabling flexible and direct modulation of action trajectory characteristics.

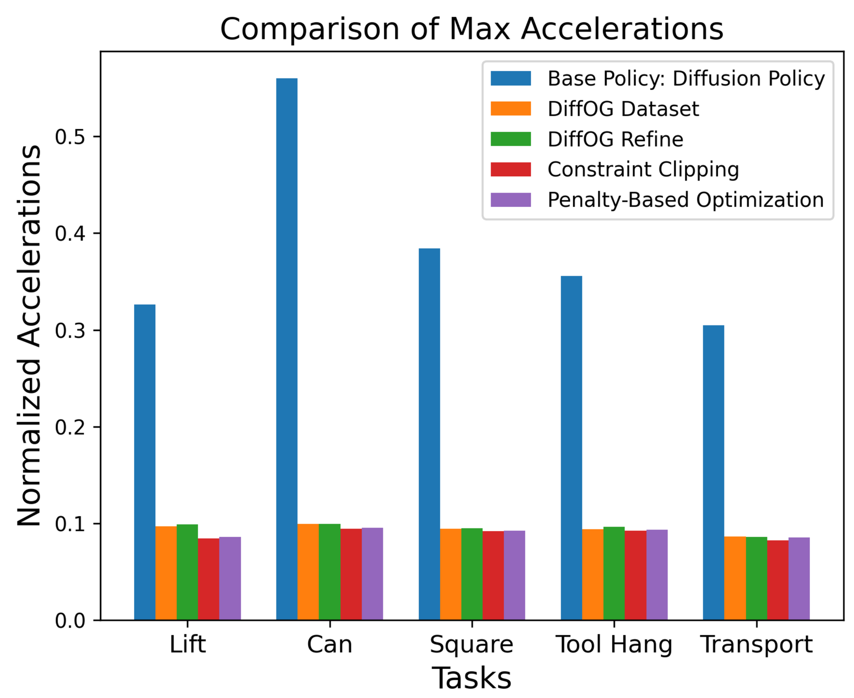


Benchmark Results

Benchmark on Visual Imitation learning

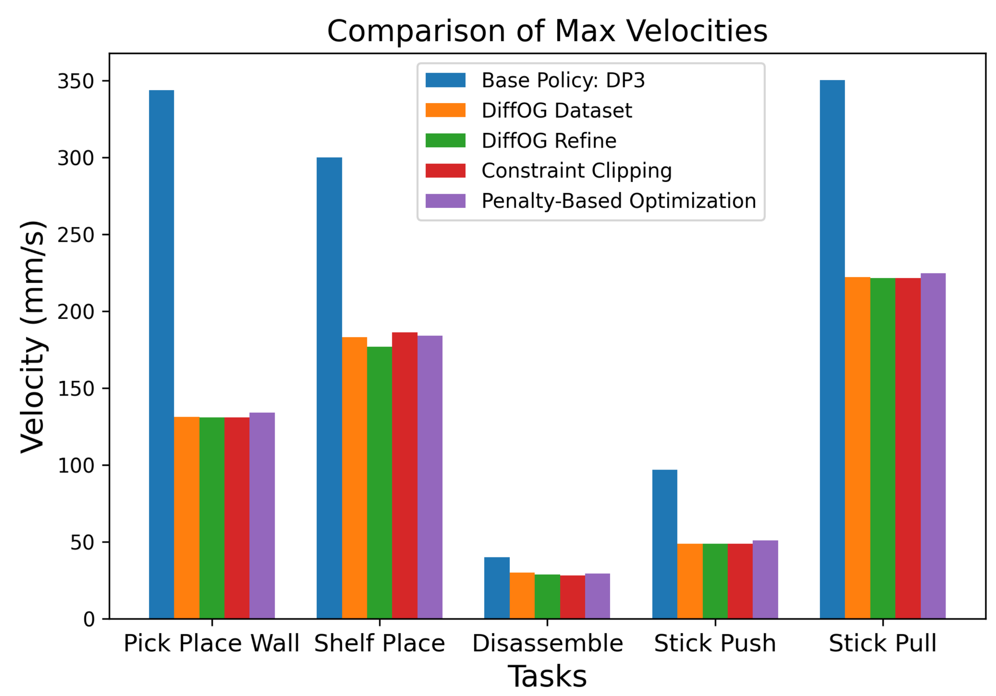
	Lift	Can	Square	Tool Hang	Transport	Push-T
Base Policy: Diffusion Policy	1.00 ± 0.00	0.98 ± 0.01	0.91 ± 0.01	0.83 ± 0.03	0.91 ± 0.01	0.84 ± 0.03
DiffOG Dataset (Ours)	$\mathbf{1.00 \pm 0.00}$	$\mathbf{0.98 \pm 0.01}$	0.87 ± 0.02	$\mathbf{0.82 \pm 0.02}$	0.89 ± 0.02	$\mathbf{0.83 \pm 0.02}$
DiffOG Refine (Ours)	$\mathbf{1.00 \pm 0.00}$	$\mathbf{0.98 \pm 0.01}$	$\mathbf{0.90 \pm 0.01}$	0.81 ± 0.02	$\mathbf{0.91 \pm 0.03}$	0.80 ± 0.02
Constraint Clipping	$\mathbf{1.00 \pm 0.00}$	0.93 ± 0.04	0.78 ± 0.03	0.59 ± 0.03	0.82 ± 0.02	0.82 ± 0.02
Penalty-Based Optimization	$\mathbf{1.00 \pm 0.00}$	0.96 ± 0.02	0.76 ± 0.03	0.79 ± 0.01	0.81 ± 0.02	0.81 ± 0.04

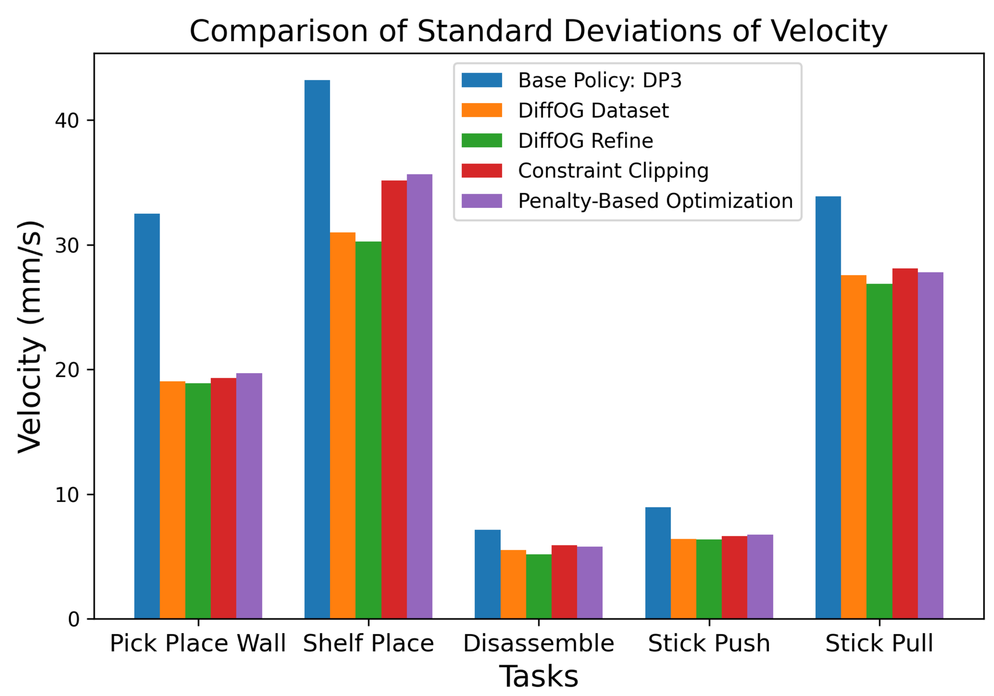




Benchmark on 3D Imitation learning

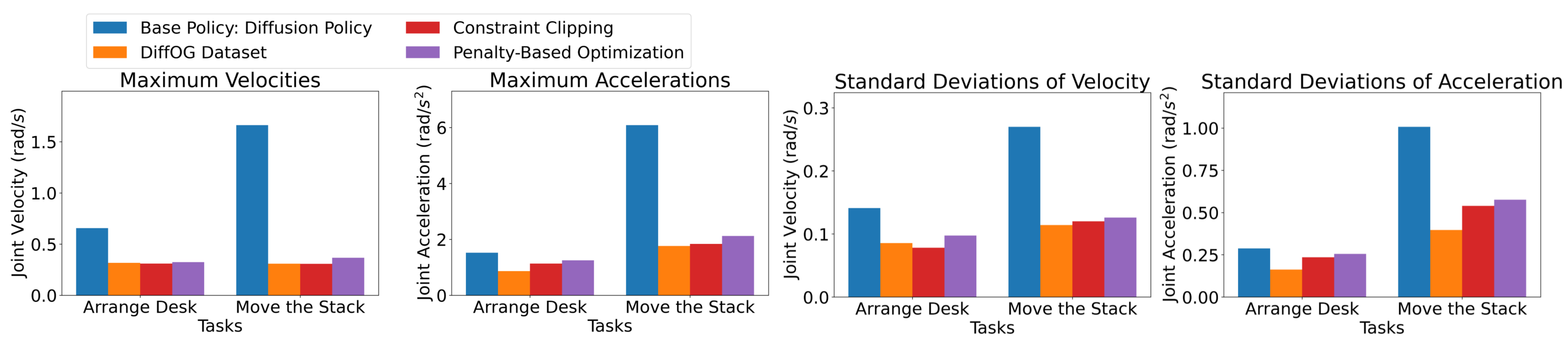
	Pick Place Wall	Shelf Place	Disassemble	Stick Push	Stick Pull
Base Policy: DP3	0.98 ± 0.01	0.77 ± 0.04	0.87 ± 0.02	1.00 ± 0.00	0.70 ± 0.02
DiffOG Dataset (Ours)	$\mathbf{0.98 \pm 0.01}$	$\mathbf{0.73 \pm 0.01}$	0.89 ± 0.05	$\mathbf{1.00 \pm 0.00}$	$\mathbf{0.70 \pm 0.03}$
DiffOG Refine (Ours)	$\mathbf{0.98 \pm 0.01}$	0.72 ± 0.02	$\mathbf{0.90 \pm 0.03}$	$\mathbf{1.00 \pm 0.00}$	0.68 ± 0.02
Constraint Clipping	$\mathbf{0.98 \pm 0.01}$	0.70 ± 0.01	0.86 ± 0.05	$\mathbf{1.00 \pm 0.00}$	0.69 ± 0.02
Penalty-Based Optimization	$\mathbf{0.98 \pm 0.01}$	0.67 ± 0.03	0.86 ± 0.06	$\mathbf{1.00 \pm 0.00}$	$\mathbf{0.70 \pm 0.03}$





Benchmark on Real-world Tasks

	Base Policy: Diffusion Policy	DiffOG Dataset	Constraint Clipping	Penalty-Based Optimizaiton
Arrange Desk	**9/15**	**9/15**	2/15	4/15
Move the Stack	3/15	**10/15**	6/15	7/15



We benchmarked DiffOG and several baselines across 13 tasks. DiffOG consistently improves action trajectories, making them smoother and more constrained, while preserving alignment with the original demonstration distribution, thereby preventing policy performance degradation.

Check out the project website!