## NeurIPS Rebuttal Experiments

	Expert Policy	Amos et al.	LSTM policy	DiffTOP(ours)
Pendulum w/o damping	13.126	$13.576 \pm 0.012$	$15.962 \pm 0.164$	$14.603 \pm 0.190$
Pendulum with dampling	10.132	$14.874 \pm 0.600$	$12.098 \pm 0.031$	$10.644 \pm 0.029$

Table 1: Cost of different algorithms on the Pendulum swingup tasks from Amos et al. As in Amos et al., we test in two settings, pendulum without damping and with damping. Lower cost means the better performance. DiffTOP performs slightly worse in the no damping case but noticeably better in the damping case.



Figure 1: Training curves of Dreamer-V3 and DiffTOP + Dreamer-V3 (Ours) on 4 DM-Control tasks: quadruped-run, quadruped-walk, reacher-hard, cheetah-run. Our method achieves better or similar performance on those tasks.