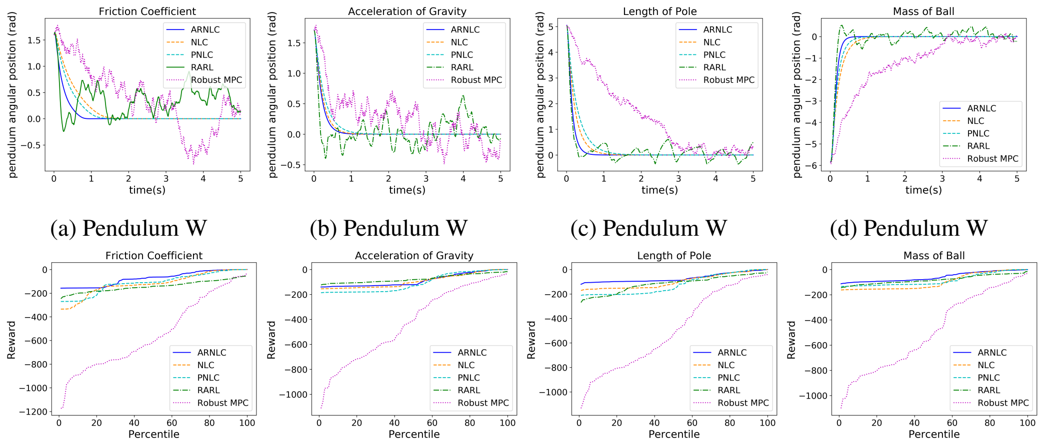
1 Evaluation under a Newly Trained Adversary

In this section, we report experimental results with evaluation under a newly trained adversary. As shown in Fig. 1, our ARNLC can achieve asymptotic stability and still holds the advantage of reaching the stability the fastest compared to the other baselines under a newly trained adversary's perturbation.



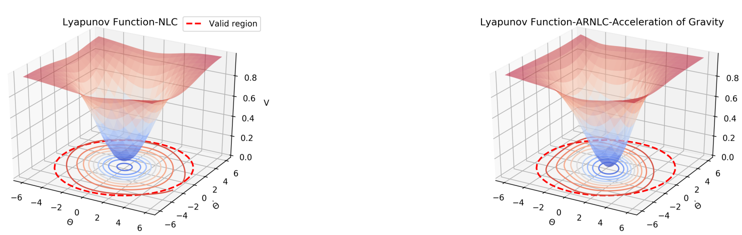
(a) Pendulum W (b) Pendulum W (c) Pendulum W (d) Pendulum W

(e) Pendulum W (f) Pendulum W (g) Pendulum W (h) Pendulum W

Figure 1: Evaluation of Pendulum under **a newly trained adversary's** worst-case (W) perturbations in testing. Figs. 1(a)-1(d) are control curves, while Figs. 1(e)-1(h) are percentile plots.

2 Comparison of Lyapunov Function

In this section, we report experimental results on the comparison of learn Lyapunov functions between original NLC and our ARNLC. As shown in Fig. 2, the Lyapunov function learned by our ARNLC and the original NLC is similar, and the contour line, which denotes the region of attraction (ROA), is also similar. Hence, our proposed ARNLC acquires a policy robust to perturbation.

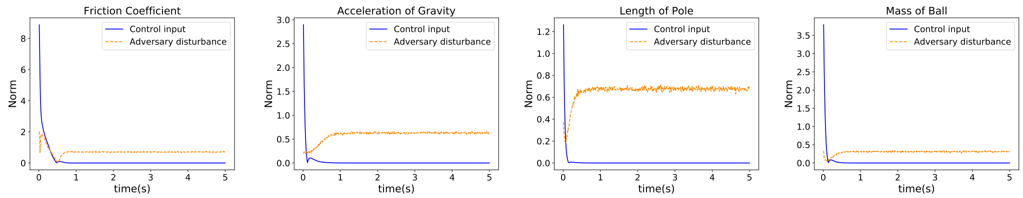


(a) NLC (b) ARNLC

Figure 2: Visualization of learned Lyapunov functions. Fig. 2(a) is for original NLC, while Fig. 2(b) is for our ARNLC.

3 Norm of Control Input and Adversary Disturbance

In this section, we report experimental results on norm of control input and adversary disturbance. As shown in Fig. 3, our controller input is in a reasonable range under the existence of the adversary perturbation.



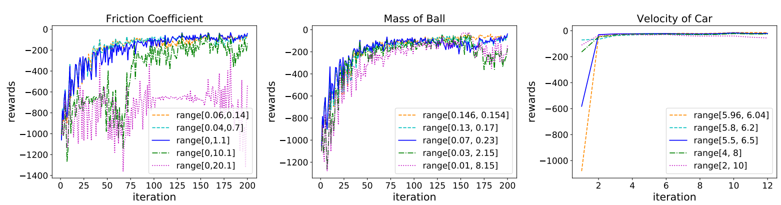
(a) Pendulum W (b) Pendulum W (c) Pendulum W (d) Pendulum W

Figure 3: Norm of control input and adversary disturbance in Pendulum task.

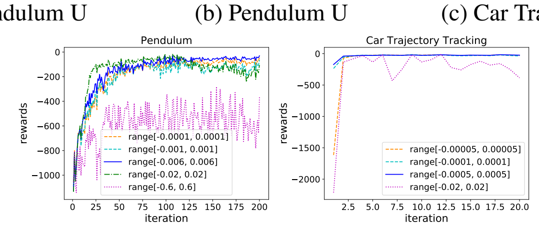
4 Comparison of Different Perturbation Ranges

In this section, we report experimental results on the comparison of different perturbation ranges applied by the adversary, to demonstrate that the perturbation range we choose can guarantee a stable training process.

We carry out experiments on both conditions of known and unknown system dynamics, respectively. The training rewards are shown in Fig. 4, where Fig. 4(a)- Fig. 4(c) are for known system dynamics and Fig. 4(d)- Fig. 4(e) are for unknown system dynamics. And control curves under uniform perturbations in testing are shown in Fig. 5, where Fig. 5(a)- Fig. 5(c) are for known system dynamics and Fig. 5(d)- Fig. 5(h) are for unknown system dynamics. We observe that a larger perturbation range applied by the adversary does result in a more unstable training process, and thus the failure in achieving asymptotic stability in most testing scenarios. Though a smaller perturbation range can guarantee a more stable training process, the performance against perturbations in testing is worse than training with larger perturbation ranges. The perturbation range plotted in blue is the perturbation range that we choose in our experiments. We tune and choose this perturbation range, which can strike the trade-off between stable training and robust controlling, i.e., it is with a stable training process as well as the ability to reach the stability with the fastest speed in the face of perturbations.

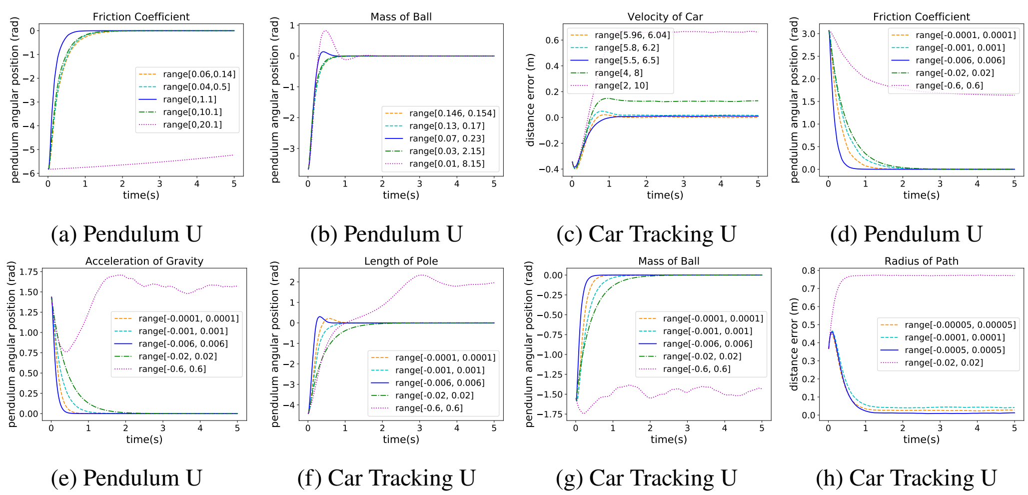


(a) Pendulum U (b) Pendulum U (c) Car Tracking U



(d) Pendulum U (e) Car Tracking U

Figure 4: Training rewards of Pendulum and Car Trajectory Tracking with different perturbation ranges applied by the adversary. Figs. 4(a)-4(c) are for known system dynamics, while Figs. 4(d)-4(e) are for unknown system dynamics.



(a) Pendulum U (b) Pendulum U (c) Car Tracking U (d) Pendulum U

(e) Pendulum U (f) Car Tracking U (g) Car Tracking U (h) Car Tracking U

Figure 5: Control curves of Pendulum and Car Trajectory Tracking under uniform (U) perturbations in testing. Figs. 5(a)-5(c) are for known system dynamics, while Figs. 5(d)-5(h) are for unknown system dynamics.