Q2: DAgger

Numerical Results Summary

Environment	Expert	BC	DAgger (Final)	Improvement
Ant-v4	4681.89	4699.91	4742.55 ± 5.59	+0.9%
HalfCheetah-v4	4034.80	4008.57	4055.00 ± 45.97	+1.2%
Hopper-v4	3717.51	1227.66	3717.14 ± 10.78	+202.8%
Walker2d-v4	5383.31	2629.40	5284.83 ± 73.11	+101.0%

Table 1: Summary of DAgger results compared to Expert and BC baselines. DAgger values show mean ± std across seeds. Improvement shows percentage change from BC baseline.

Learning Curves

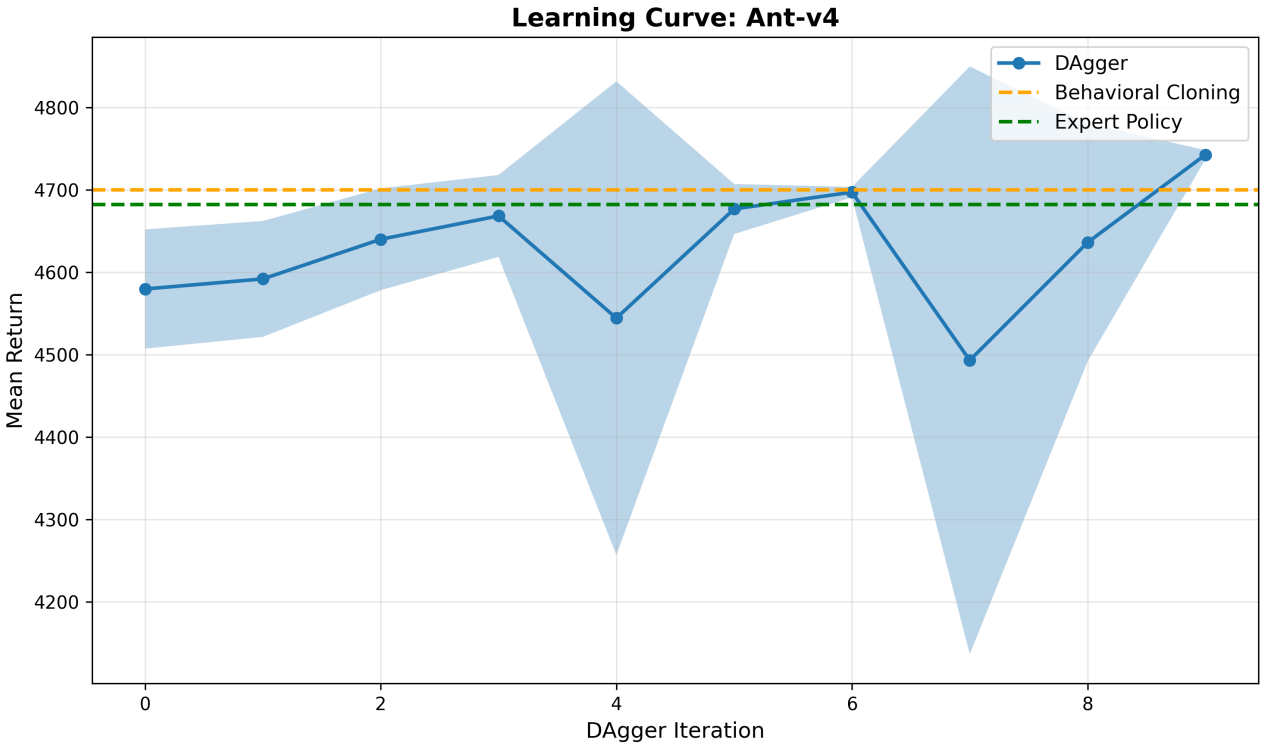

Figure 1: **DAgger Learning Curve for Ant-v4.** The plot shows mean return over 3 seeds with error bars (standard deviation). Network architecture: 2-layer MLP with 64 hidden units per layer. Training: 1000 gradient steps per iteration, batch size 1000. Final DAgger performance: 4742.55 ± 5.59. BC baseline: 4699.91. Expert performance: 4681.89.

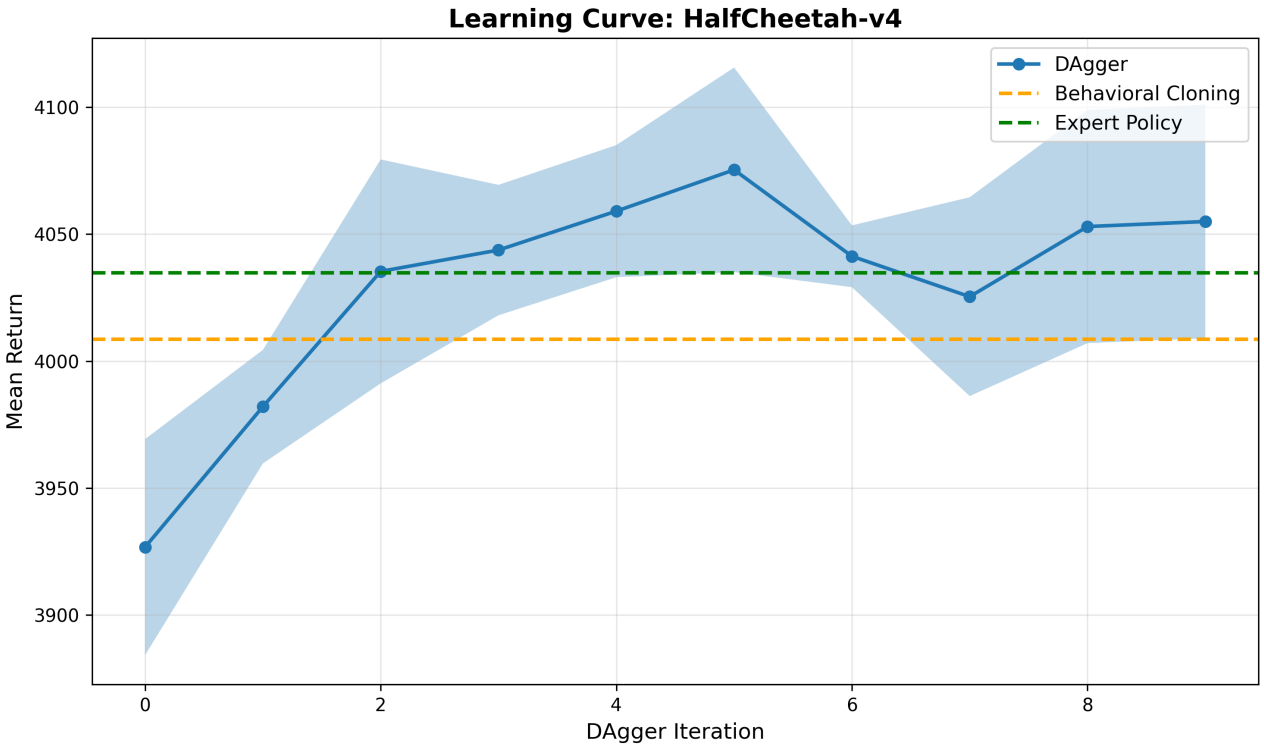

Figure 2: **DAgger Learning Curve for HalfCheetah-v4.** Final DAgger performance: 4055.00 ± 45.97. BC baseline: 4008.57. Expert performance: 4034.80.

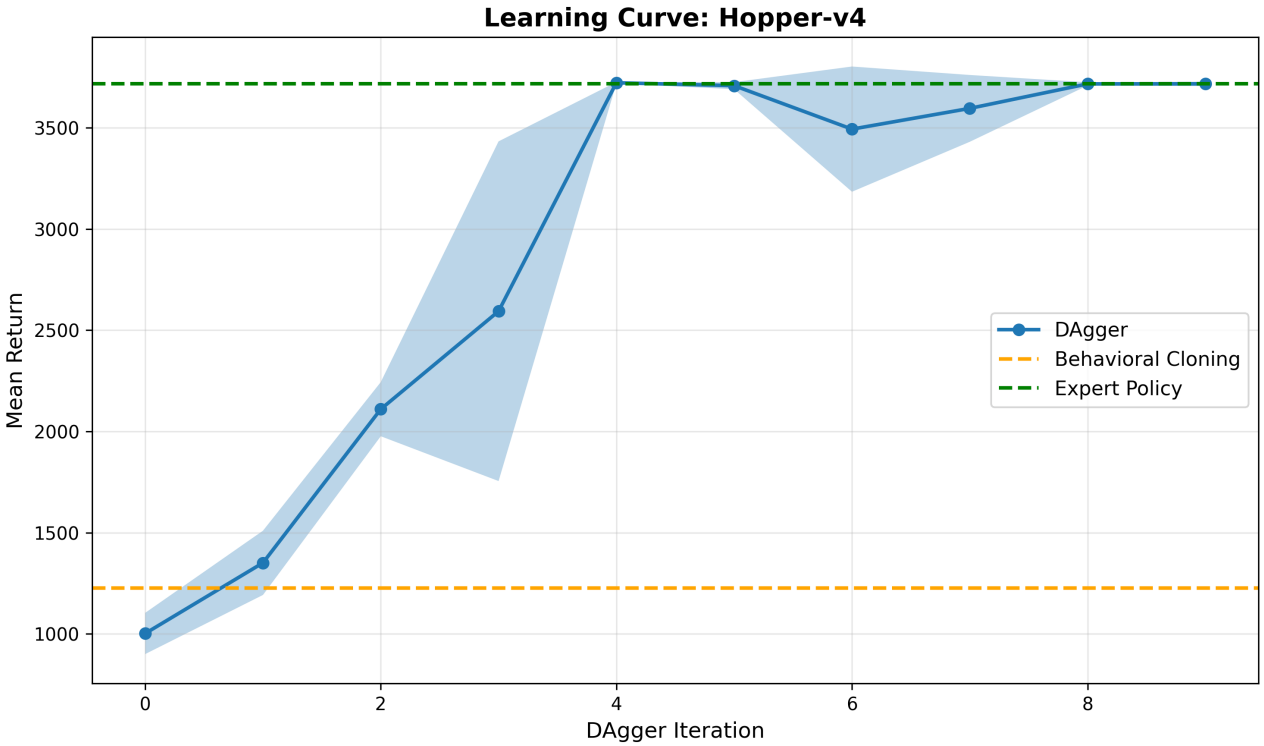

Figure 3: **DAgger Learning Curve for Hopper-v4.** Final DAgger performance: 3717.14 ± 10.78. BC baseline: 1227.66. Expert performance: 3717.51.

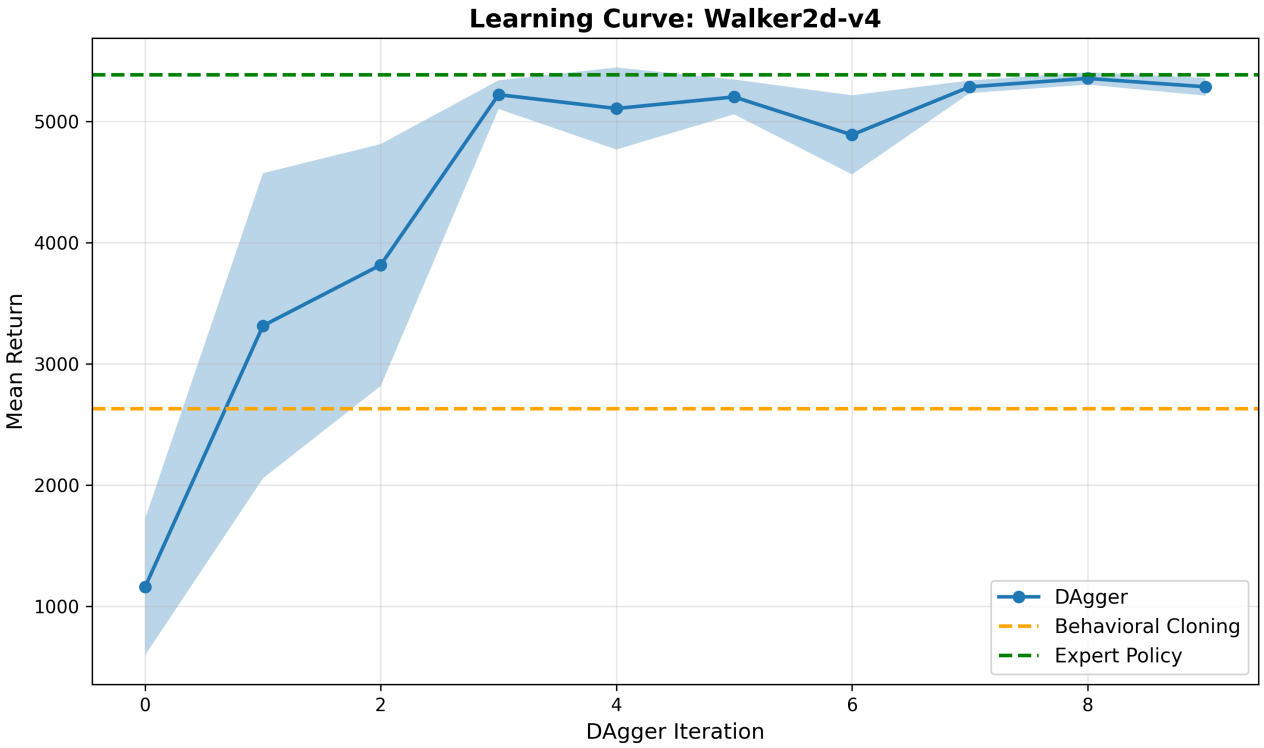

Figure 4: **DAgger Learning Curve for Walker2d-v4.** Final DAgger performance: 5284.83 ± 73.11. BC baseline: 2629.40. Expert performance: 5383.31.