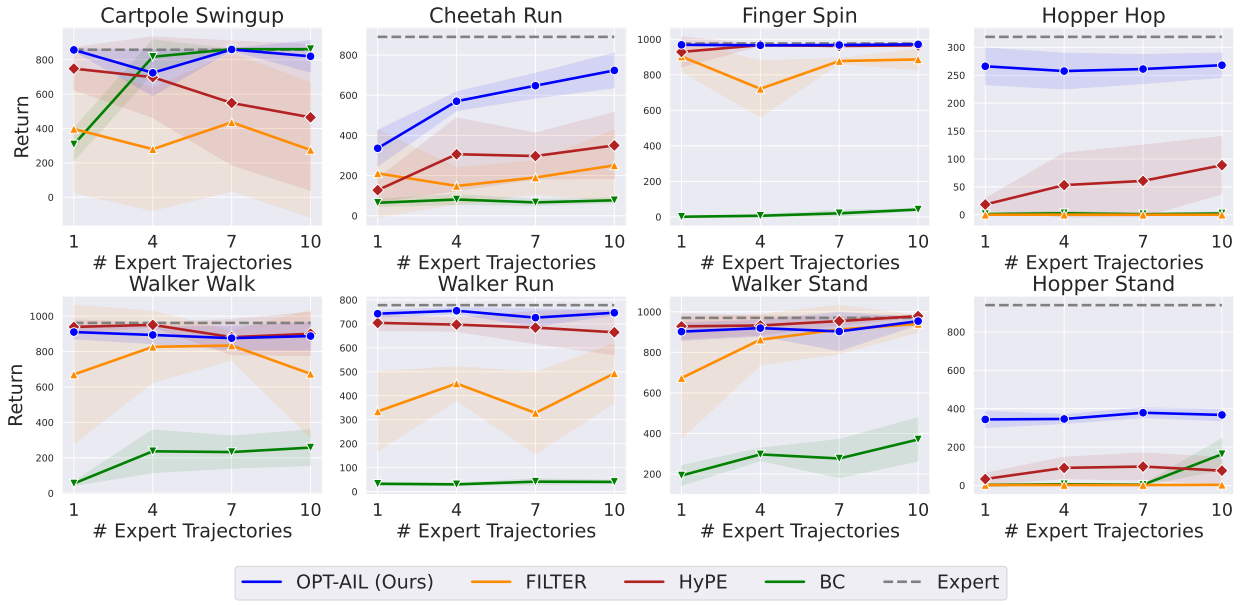

Figure 1: Overall performance of BC, FILTER, HyPE and OPT-AIL on 8 DMControl tasks over 5 random seeds following 500k interactions with the environment. Here the x-axis is the number of expert trajectories and the y-axis is the return. The solid lines are the mean of results while the shaded region corresponds to the standard deviation over 5 random seeds. We find that OPT-AIL consistently matches or exceeds the overall performance of FILTER and HyPE.

Figure 2: Learning curves of BC, FILTER, HyPE and OPT-AIL on 8 DMControl tasks over 5 random seeds using 4 expert trajectories. Here the x-axis is the number of environment interactions and the y-axis is the return. The solid lines are the mean of results while the shaded region corresponds to the standard deviation over 5 random seeds. We find that OPT-AIL can achieve near-expert performance with fewer environment interactions compared with FILTER and HyPE.