Conservative Q-Learning (CQL) Training Dynamics Critic Loss Actor Loss 1e1 7.0 6.5 -56.0 **Nale** 5.5 5.0 Value -10-154.5 -20 4.0 3.5 -255000 10000 12500 20000 2500 5000 7500 10000 12500 2500 7500 15000 17500 15000 17500 20000 **Training Steps Training Steps SAC Temperature (Alpha) CQL Conservative Term** 1.0 3.0 2.5 0.8 2.0 0.6 Value **Najne** 1.5 0.4 1.0 0.5 0.2 0.0 10000 12500 10000 2500 5000 7500 15000 17500 20000 0 2500 5000 7500 12500 15000 17500 20000

Training Steps

Training Steps