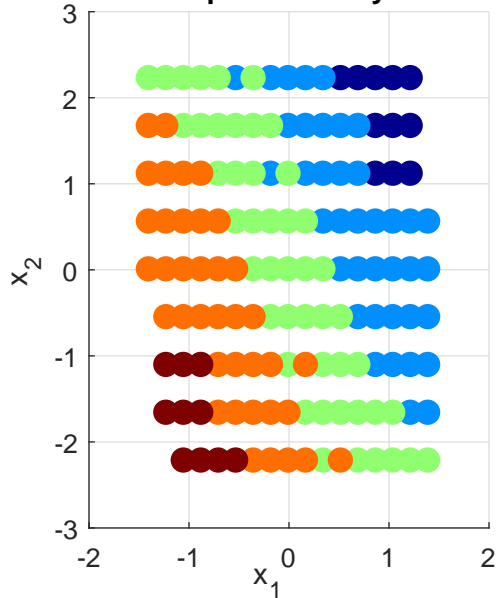
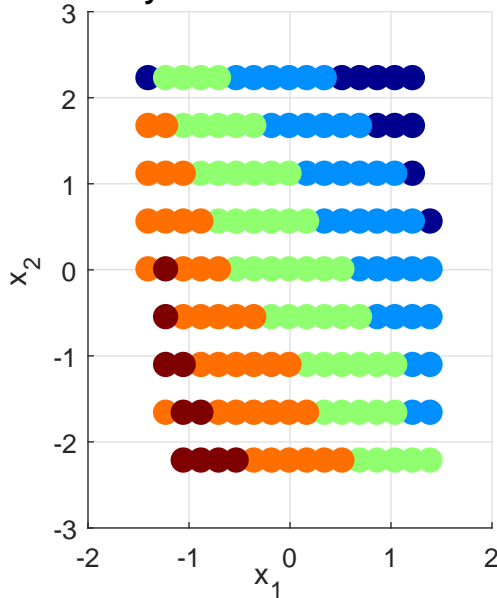


Optimal Policy



Est. Policy with Incremental Q-learning



Est. Policy w/o Exploration

