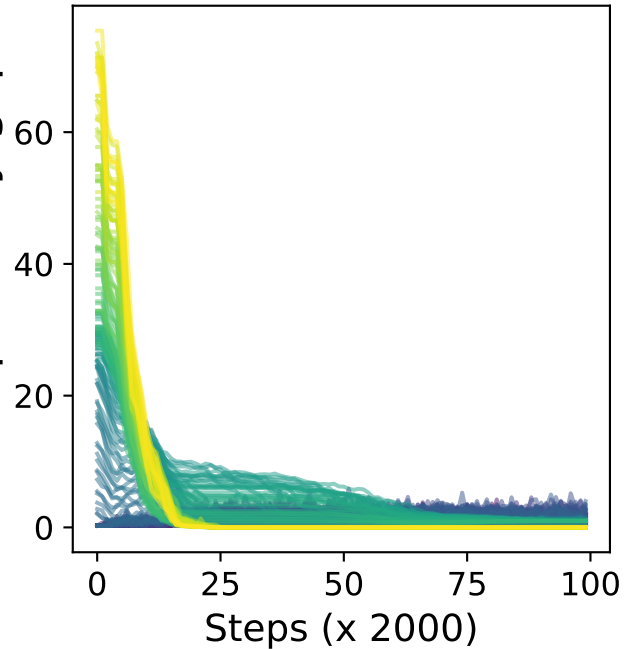
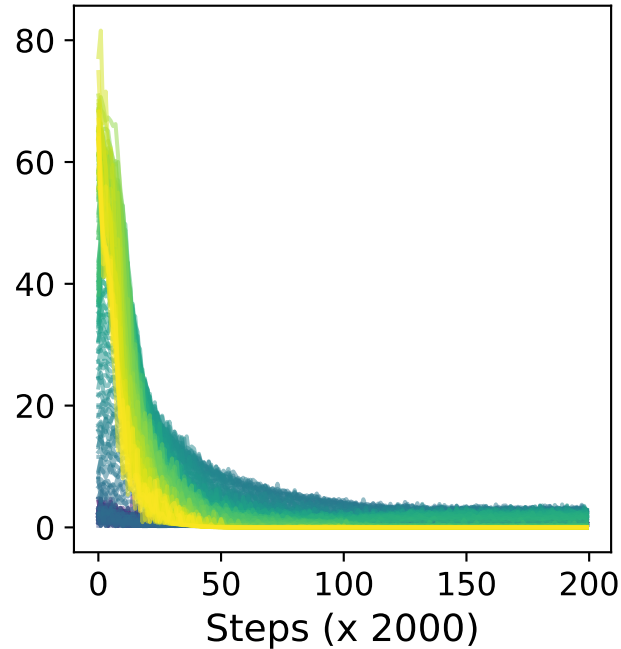


Sub-optimality gap

Q-Learning



Sarsa



Policy Gradient

