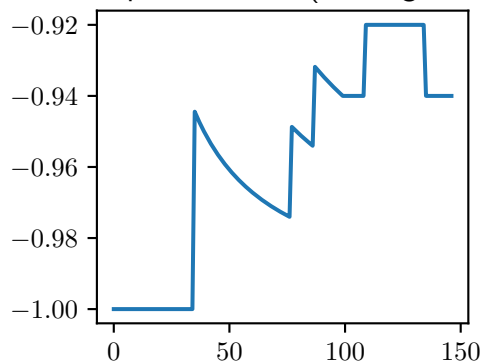
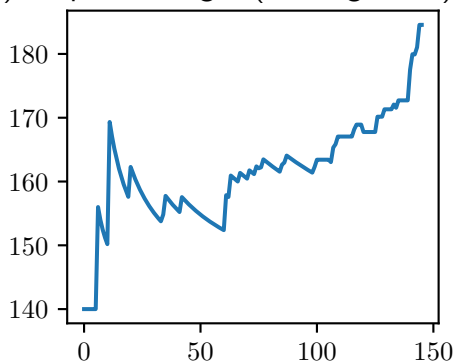


DQN training progress on pong task

100-episode return (running mean)



Episode length (running mean)



— indiv. run

Behavior entropy (running mean) Sample loss (RMSE) (running mean)

