DQN (sequential samples) Episode regret Episode regret (smoothed) 15.0 12.5 10.0 Mean regret 7.5 5.0 2.5 0.0

400

Episode number

600

800

1000

200

0