episode_55_steps_32000 Total reward explorative Total steps explorative KL Div w.r.t. optimal allocation - generative 10^{4} 60 1.2 -50 1.0 8.0 40 10^{3} 30 0.6 0.4 20 10^{2} 0.2 10 0.0 50 10 20 30 40 10 20 30 40 50 5000 10000 15000 20000 25000 30000 0 0 Total reward greedy Total steps greedy KL Div w.r.t. optimal allocation - navig. constraints 400 1.2 10^{1} 350 1.0 300 8.0 10^{0} 250 0.6 -200 150 0.4 10^{-1} 100 0.2 -50 10^{-2} 10 20 30 40 50 10 20 30 40 50 5000 10000 15000 20000 25000 30000 0

Episode

Iteration

Episode