episode_3_steps_500 Total reward explorative Total steps explorative KL Div w.r.t. optimal allocation - generative 10^{4} 1.2 -60 -1.0 -50 8.0 40 10^{3} 0.6 30 0.4 20 10^{2} 0.2 10 0.0 0.00 0.25 0.50 0.75 1.00 1.25 1.50 1.75 2.00 0.00 0.25 0.50 0.75 1.00 1.25 1.50 1.75 2.00 100 200 300 400 500 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} 0.0 -0.04-0.020.00 0.02 0.04 -0.04-0.020.00 0.02 0.04 100 200 300 400 500 0

Episode

Iteration

Episode