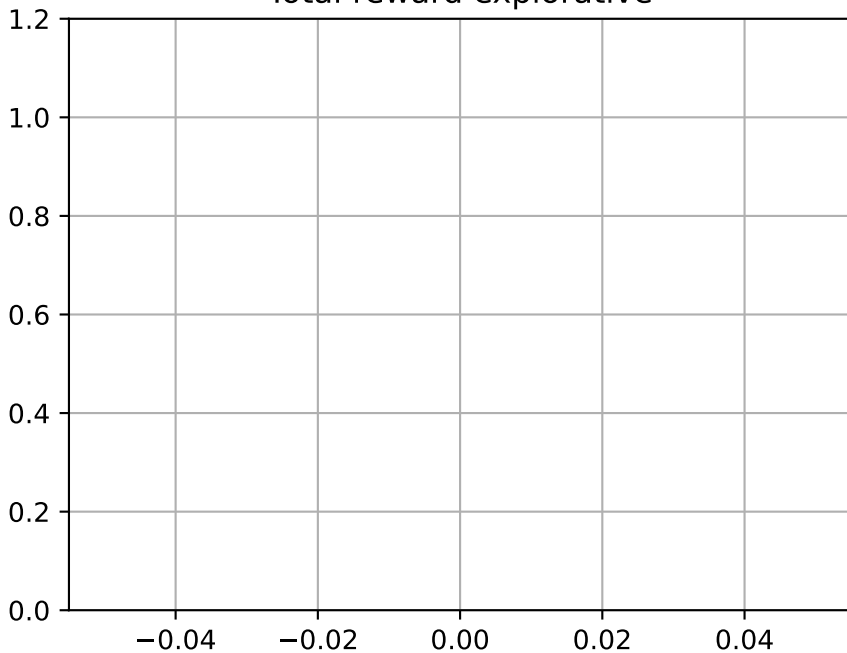
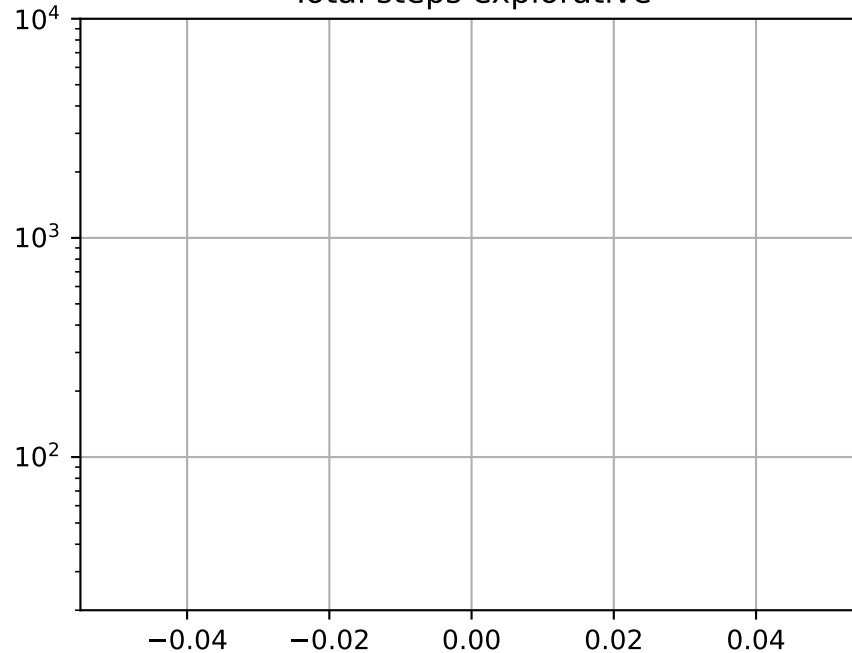


episode_1_steps_3500

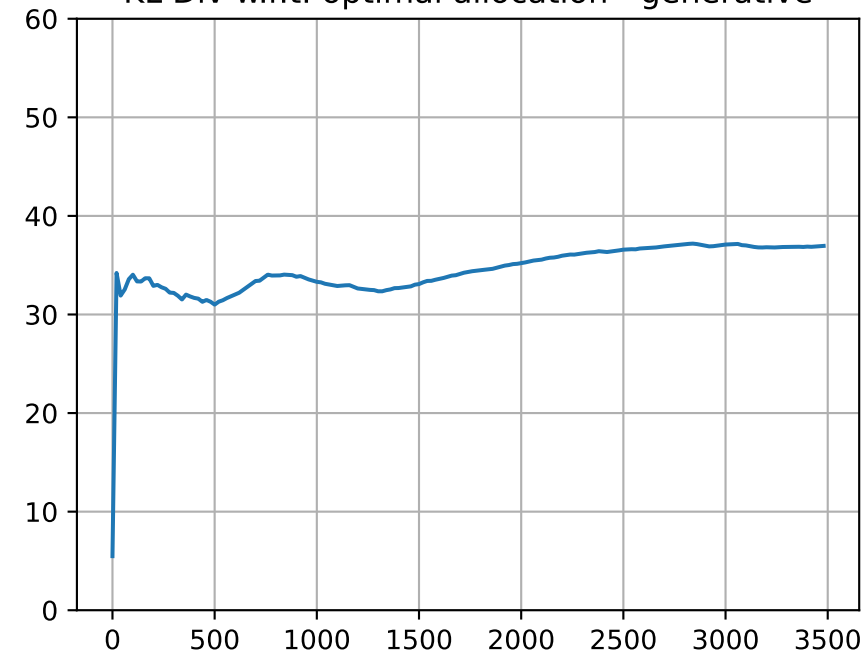
Total reward explorative



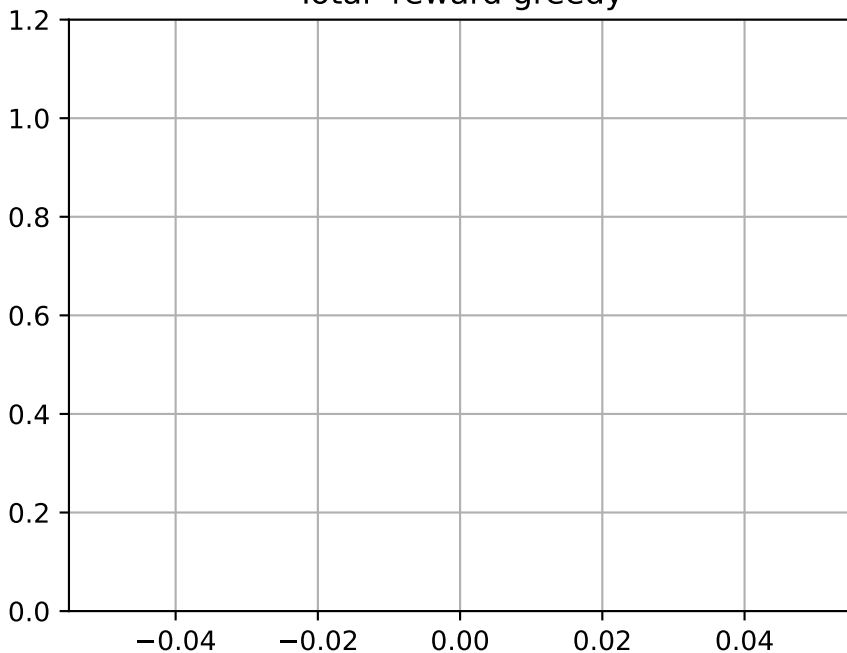
Total steps explorative



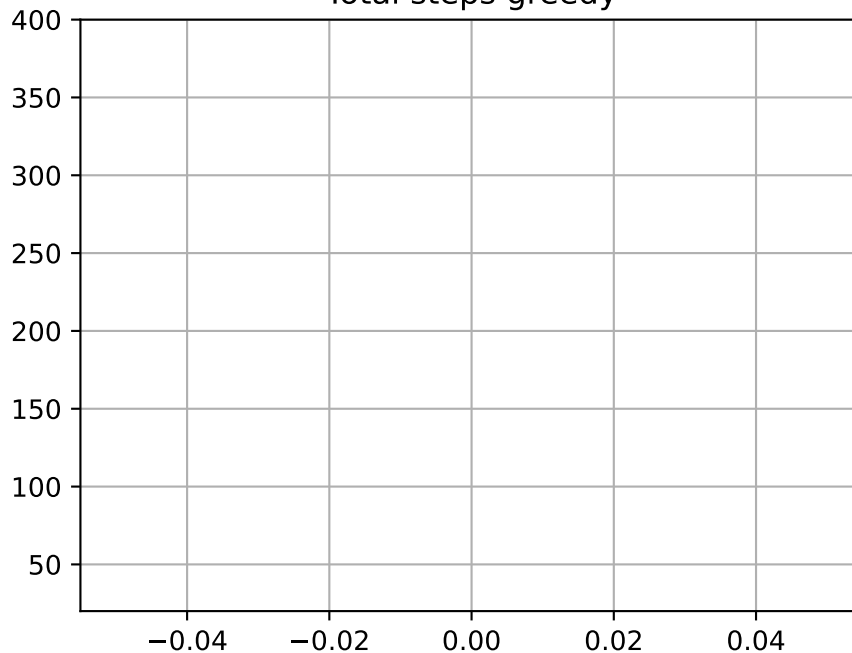
KL Div w.r.t. optimal allocation - generative



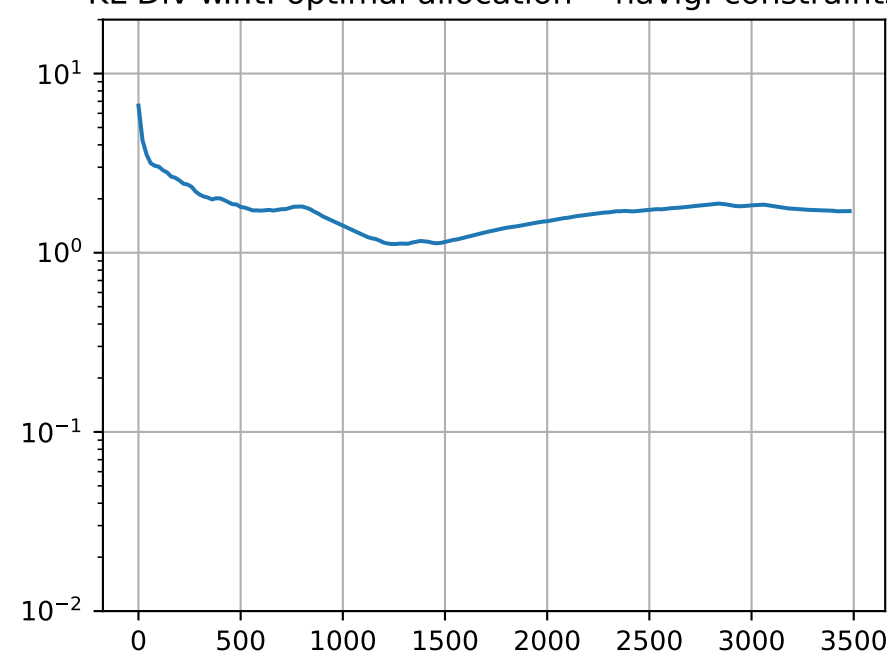
Total reward greedy



Total steps greedy



KL Div w.r.t. optimal allocation - navig. constraints



Episode

Episode

Iteration